

OPTIMAL NETWORKS

A. O. Ivanov and A. A. Tuzhilin

The aim of this mini-course is to give an introduction in Optimal Networks theory. Optimal networks appear as solutions of the following natural problem: How to connect a finite set of points in a metric space in an optimal way? We cover three most natural types of optimal connection: spanning trees (connection without additional road forks), shortest trees and locally shortest trees, and minimal fillings.

1 Introduction: Optimal Connection

This mini-course was given in the First Yaroslavl Summer School on Discrete and Computational Geometry in August 2012, organized by International Delaunay Laboratory “Discrete and Computational Geometry” of Demidov Yaroslavl State University. We are very thankful to the organizers for a possibility to give lectures there and to publish this notes, and also for their warm hospitality during the Summer School. The real course consisted of three 1 hour lectures, but the division of these notes into sections is independent on the lectures structure. The video of the lectures can be found in the site of the Laboratory (<http://dcglab.uniyar.ac.ru>). The main reference is our books [1] and [2], and the paper [3] for Section 5.

Our subject is optimal connection problems, a very popular and important kind of geometrical optimization problems. We all seek what is better, so optimization problems attract specialists during centuries. Geometrical optimization problems related to investigation of critical points of geometrical functionals, such as length, volume, energy, etc. The main example for us is the length functional, and the corresponding optimization problem consists in finding of length minimal connections.

1.1 Connecting Two Points

If we have to points A and B in the Euclidean plane \mathbb{R}^2 , then, as we know from the elementary school, the shortest curve joining A and B is unique and coincides with the straight segment AB , so optimal connection problem is trivial in this case. But if we change the way of distance measuring and consider, for example, so-called Manhattan plane, i.e. the plane \mathbb{R}^2 with fixed standard coordinates (x, y) and the distance function $\rho_1(A, B) = |a_1 - b_1| + |a_2 - b_2|$, where $A = (a_1, a_2)$ and $B = (b_1, b_2)$, then it is not difficult to verify that in this case there are infinitely many shortest curves connecting A and B . Namely, if $0 \leq a_1 < b_1$ and $0 \leq a_2 < b_2$, then any monotonic curve $\gamma(t) = (x(t), y(t))$, $t \in [0, 1]$, $\gamma(0) = A$, $\gamma(1) = B$, where functions $x(t)$ and $y(t)$ are monotonic, are the shortest, see Figure 1, left. Another new effect that can be observed in this example is as follows. In the Euclidean plane a curve such that each its sufficiently small piece is a shortest curve joining its ends (so-called *locally shortest curve*) is a shortest curve itself. In the Manhattan plane it is not so. The length of a locally shortest curve having the form of the letter Π , see Figure 1, right, can be evidently decreased.

Similar effects can be observed in the surface of standard sphere $S^2 \subset \mathbb{R}^3$. Here the shortest curve joining a pair of points is the lesser arc of the great circle (the cross-section of the sphere by a plane passing through the origin). Two opposite points are connected

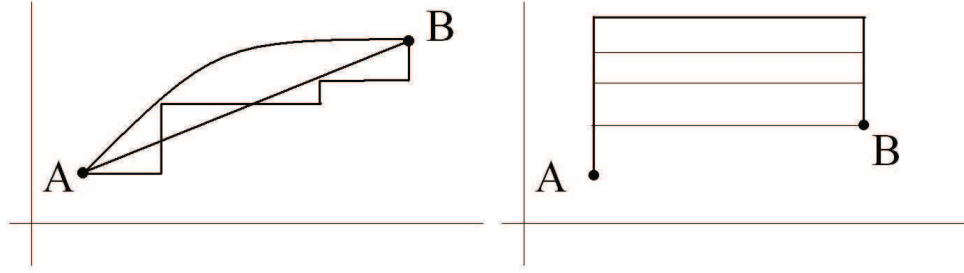


Figure 1: The shortest curves connecting a pair of points in Manhattan plane (left), and locally shortest but not shortest curve in this plane (right).

by infinitely many shortest curves, and if points A and B are not opposite, then the corresponding great circle is unique and it is partitioned into two arcs, both of them are locally shortest, one is the (unique) shortest, but the other one is not. (Really speaking, the difference with the Manhattan plane consists in the fact that for the case of the sphere any directional derivative of the length of any locally shortest arc with respect to its deformation preserving the ends is equal to zero).

Exercise 1.1 For a pair of points on the surface of the cube describe shortest and locally shortest curves. Find out an infinite family of locally shortest curves having pairwise distinct lengths.

1.2 Connecting Many Points: Possible Approaches

Let us consider general situation, when we are given with a finite set $M = \{A_1, \dots, A_n\}$ of points in a metric space (X, ρ) , and we want to connect them in some optimal way in the sense of the total length of the connection. We are working under assumption that we already know how to connect pairs of points in (X, ρ) , therefore we need just to organize the set of shortest curves in appropriate way. There are several natural statements of the problem, and we list here the most popular ones.

1.2.1 No Additional Forks Case: Spanning Trees

We do not allow additional forks, that is, we can switch between the shortest segments at the points from M only. As a result, we obtain a particular case of Graph Theory problem about minimal spanning trees in a connected weighted graph. We recall only necessary concepts of Graph Theory, the detail can be found, for example in [4].

Recall that a (*simple*) *graph* can be considered as a pair $G = (V, E)$, consisting of a finite set $V = \{v_1, \dots, v_n\}$ of *vertices* and a finite set $E = \{e_1, \dots, e_m\}$ of *edges*, where each edge e_i is a two-element subset of V . If $e = \{v, v'\}$, then we say that v and v' are *neighboring*, edge e *joins* or *connects* them, the edge e and each of the vertices v and v' are *incident*. The number of vertices neighboring to a vertex v is called the *degree* of v and is denoted by $\deg v$. A graph $H = (V_H, E_H)$ is said to be a *subgraph* of a graph $G = (V_G, E_G)$, if $V_H \subset V_G$ and $E_H \subset E_G$. The subgraph H is called *spanning*, if $V_H = V_G$.

A *path* γ in a graph G is a sequence $v_{i_1}, e_{i_1}, v_{i_2}, \dots, e_{i_k}, v_{i_{k+1}}$ of its vertices and edges such that each edge e_{i_s} connects vertices v_{i_s} and $v_{i_{s+1}}$. We also say that the path γ connects the vertices v_{i_1} and $v_{i_{k+1}}$ which are said to be *ending vertices* of the path. A

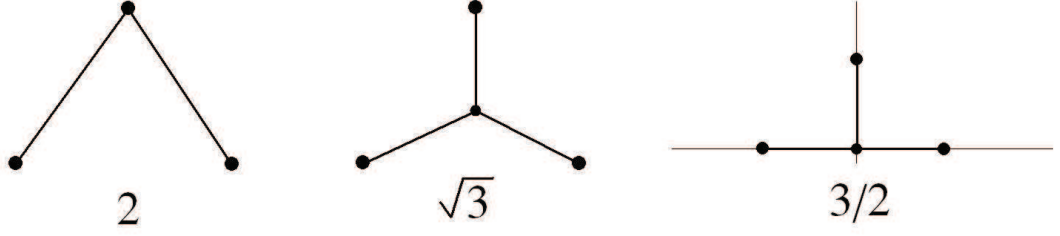


Figure 2: Minimal spanning tree (left), shortest tree (center), and minimal filling, connecting the vertex set of regular triangle in Euclidean plane.

path is said to be *cyclic*, if its ending vertices coincide with each other. A cyclic path with pairwise distinct edges is referred as a *simple cycle*. A graph without simple cycles is said to be *acyclic*. A graph is said to be *connected*, if any its two vertices can be connected by a path. An acyclic connected graph is called a *tree*.

If we are given with a function $\omega: E \rightarrow \mathbb{R}$ on the edge set of a graph G , then the pair (G, ω) is referred as a *weighted graph*. For any subgraph $H = (V_H, E_H)$ of a weighted graph $(G = (V_G, E_G), \omega)$ the value $\omega(H) = \sum_{e \in E_H} \omega(e)$ is called the *weight of H* . Similarly, for any path $\gamma = v_{i_1}, e_{i_1}, v_{i_2}, \dots, e_{i_k}, v_{i_{k+1}}$ the value $\omega(\gamma) = \sum_{s=1}^k \omega(e_{i_s})$ is called the *weight of γ* .

For a weighted connected graph $(G = (V_G, E_G), \omega)$ with positive weight function ω , a spanning connected subgraph of minimal possible weight is called *minimal spanning tree*. The positivity of ω implies that such subgraph is *acyclic*, i.e. it is a tree indeed. The weight of any minimal spanning tree for (G, ω) is denoted by $\text{mst}(G, \omega)$.

Optimal connection problem without additional forks can be considered as minimal spanning tree problem for a special graph. Let $M = \{A_1, \dots, A_n\}$ be a finite set of points in a metric space (X, ρ) as above. Consider the complete graph $K(M)$ with vertex set M and edge set consisting of all two-element subsets of M . In other words, any two vertices A_i and A_j are connected by an edge in $K(M)$. By $A_i A_j$ we denote the corresponding edge. The number of edges in $K(M)$ is, evidently, $n(n-1)/2$. We define the positive weight function $\omega_\rho(A_i A_j) = \rho(A_i, A_j)$. Then any minimal spanning tree T in $K(M)$ can be considered as a set of shortest curves in (X, ρ) joining corresponding points and forming a network in X connecting M without additional forks in an optimal way, i.e. with the least possible length. Such a network is called a *minimal spanning tree for M in (X, ρ)* . Its total weight $\omega_\rho(T)$ is called *length* and is denoted by $\text{mst}_X(M)$. In Section 2 we speak about minimal spanning trees in more details.

1.2.2 Shortest tree: Fermat–Steiner Problem

But already P. Fermat and C. F. Gauss understood that additional forks can be profitable, i.e. can give an opportunity to decrease the length of optimal connection. For example, see Figure 2, if we consider the vertex set $M = \{A_1, A_2, A_3\}$ of a regular triangle with side 1 in the Euclidean plane, then the corresponding graph $K(M)$ consists of three edges of the same weight 1 and each minimal spanning tree consists of two edges, so $\text{mst}_{\mathbb{R}^2}(M) = 2$. But if we add the center T of the triangle and consider the network consisting of three straight segments $A_1 T, A_2 T, A_3 T$, then its length is equal to $3 \frac{2}{3} \frac{\sqrt{3}}{2} = \sqrt{3} < 2$, so it is shorter than the minimal spanning tree.

This reasoning leads to the following general definition. Let $M = \{A_1, \dots, A_n\}$ be

a finite set of points in a metric space (X, ρ) as above. Consider a larger finite set N , $M \subset N \subset X$, and a minimal spanning tree for N in X . Then this tree contains M as a subset of its vertex set N , but also may contain some other additional vertices—forks. Such additional vertices are referred as *Steiner points*. Further, we define a value $\text{smt}_X(M) = \inf_{N: M \subset N \subset X} \text{mst}_X(N)$ and call it the *length of shortest tree connecting M* or of *Steiner minimal tree for M* . If this infimum attains at some set N , then each minimal spanning tree for this N is called a *shortest tree* or a *Steiner minimal tree* connecting M . Famous Steiner problem is the problem of finding a shortest tree for a given finite subset of a metric space. We will speak about Steiner problem in more details in Section 3. The shortest tree for the vertex set of a regular triangle in the Euclidean plane is depicted in Figure 2.

1.2.3 Minimizing over Different Ambient Spaces: Minimal Fillings

Shortest trees give the least possible length of connecting network for a given finite set in a fixed ambient space. But sometimes it is possible to decrease the length of connection by choosing another ambient space. Let $M = \{A_1, \dots, A_n\}$ be a finite set of points in a metric space (X, ρ) as above, and consider M as a finite metric space with the distance function ρ_M obtained as the restriction of the distance function ρ . Consider an isometric embedding $\varphi: (M, \rho_M) \rightarrow (Y, \rho_Y)$ of this finite metric space (M, ρ_M) into a (compact) metric space (Y, ρ_Y) and consider the value $\text{smt}_Y(\varphi(M))$. It could be less than $\text{smt}_X(M)$. For example, the vertex set of the regular triangle with side 1 can be embedded into Manhattan plane as the set $\{(-1/2, 0), (0, 1/2), (1/2, 0)\}$, see Figure 2. Then the unique additional vertex of the shortest tree is the origin and the length of the tree is $3/2 < \sqrt{3}$. So, for a finite metric space $\mathcal{M} = (M, \rho_M)$, consider the value $\text{mf}(\mathcal{M}) = \inf_{\varphi} \text{smt}_Y(\varphi(M))$ which is referred as *weight of minimal filling* of the finite metric space \mathcal{M} . Minimal fillings can be naturally defined in terms of weighted graphs and can be considered as a generalization of Gromov’s concept of minimal fillings for Riemannian manifolds. We speak about them in more details in Section 5.

2 Minimal Spanning Trees

In this section we discuss minimal spanning trees construction in more details. As we have already mentioned above, in this case the problem can be stated in terms of Graph Theory for an arbitrary connected weighted graph. But geometrical interpretation permits to speed up the algorithms of Graph Theory.

2.1 General Case: Graph Theory Approach

We start with the Graph Theory problem of finding a minimal spanning tree in a connected weighted graph. It is not difficult to verify that direct enumeration of all possible spanning subtrees of a connected graph leads to an exponential algorithm.

To see that, recall well-known Kirchhoff theorem counting the number of spanning subtrees. If $G = (V, E)$ is a connected graph with enumerated vertex set $V = \{v_1, \dots, v_n\}$, then its *Kirchhoff matrix* is defined as $(n \times n)$ -matrix $B_G = (b_{ij})$ with elements

$$b_{ij} = \begin{cases} \deg v_i & \text{if } i = j, \\ -1 & \text{if } \{v_i, v_j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Then the following result based on elementary Graph Theory and Binet–Cauchy formula for determinant calculation is valid, see proof, for example in [4].

Theorem 2.1 (Kirchhoff) *For a connected graph G with $n \geq 2$ vertices, the number of spanning subtrees is equal to the algebraic complement of any element of the Kirchhoff matrix B_G .*

Example. Let $G = K_n$ be the complete graph with n vertices. Then its Kirchhoff matrix has the following form:

$$B_{K_n} = \begin{pmatrix} n-1 & -1 & -1 & \cdots & -1 \\ -1 & n-1 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & n-1 \end{pmatrix}.$$

The algebraic complement of the element b_{nn} is equal to

$$\begin{vmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{vmatrix} = \begin{vmatrix} 1 & \cdots & 1 \\ -1 & \cdots & -1 \\ \vdots & \ddots & \vdots \\ -1 & \cdots & n-1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{vmatrix} = n^{n-2},$$

where the first equality is obtained by change of the first row by the sum of all the rows, and the second equality is obtained by change of the i th row, $i \geq 2$, by the sum of it with the first row.

Corollary 2.2 *The complete graph with n vertices contains n^{n-2} spanning trees.*

Remark. Notice that this result is equivalent to Cayley Theorem saying that the total number of trees with n enumerated vertices is equal to n^{n-2} .

But it is a surprising fact, that there exist polynomial algorithms constructing minimal spanning trees. Several such algorithms were discovered in 1960s. We tell about Kruskal's algorithm. Similar Prim's algorithm can be found in [4].

So, we are given with a connected weighted graph $(G = (V, E), \omega)$ with positive weight function ω . At the initial step of Kruskal algorithm we construct the graph $T_0 = (V, \emptyset)$ and put $E_0 = E$. If the graph T_{i-1} and the non-empty set $E_{i-1} \subset E$, $i \leq n-1$, have been already constructed, then we choose in E_{i-1} an edge e_i of least possible weight and construct a new graph $T_i = T_{i-1} \cup e_i$ and also a new set $E_i = \{e \in E \mid e \notin T_i \text{ and } T_i \cup e \text{ is acyclic}\}$. Algorithm stops when the graph T_{n-1} is constructed¹.

Theorem 2.3 (Kruskal) *Under the above notations, the graph T_{n-1} can be constructed for any connected weighted graph $\mathcal{G} = (G, \omega)$, and moreover, T_{n-1} is a minimal spanning tree in \mathcal{G} .*

Proof. The set E_i is non-empty for all $0 \leq i \leq n-2$, because the corresponding subgraphs T_i are not connected (the graph T_i has n vertices and i edges), therefore all the graphs T_1, \dots, T_{n-1} can be constructed. Further, all these graphs are acyclic due to the construction, and T_{n-1} has n vertices and $n-1$ edges, so it is a tree.

To finish the proof it remains to show that the spanning tree $T_{n-1} \subset G$ is minimal. Since the graph \mathcal{G} has a finite number of spanning trees, a minimal spanning tree does exist. Let T be a minimal spanning tree. We show that it can be reconstructed to the tree T_{n-1} without changing the total weight, so T_{n-1} is also a minimal spanning tree.

¹ Here the operation of adding an edge e to a graph $G = (V, E)$ can be formally defined as follows: $G \cup e = (V, E \cup \{e\})$. Similarly, $G \setminus e = (V, E \setminus \{e\})$.

To do this, recall that the edges of the tree T_{n-1} are enumerated in accordance with the work of the algorithm. Denote them by e_1, \dots, e_{n-1} as above, and assume that e_k is the first one that does not belong to T . The graph $T \cup e_k$ contains a unique cycle $c \supset e_k$. This cycle c also contains an edge e not belonging to T_{n-1} (otherwise $c \subset T_{n-1}$, a contradiction). Consider the graph $T' = T \cup e_k \setminus e$. It is evidently a spanning tree in G , and therefore its weight is not less than the weight of the minimal spanning tree T , hence

$$\omega(T') = \omega(T) + \omega(e_k) - \omega(e) \geq \omega(T),$$

and thus, $\omega(e_k) \geq \omega(e)$.

On the other hand, all the edges e_1, \dots, e_{k-1} belongs to T by our assumption. Therefore, the graph $T_{k-1} \cup e$ is a subgraph of T and is acyclic, in particular. Hence, $e \in E_{k-1}$ so as $e_k \in E_{k-1}$. But the algorithm has chosen e_k , hence $\omega(e_k) \leq \omega(e)$. Thus, $\omega(e_k) = \omega(e)$, and so $\omega(T') = \omega(T)$, and therefore T' is a minimal spanning tree in \mathcal{G} . But now T' contains the edges e_1, \dots, e_k from T_{n-1} . Repeating this procedure we reconstruct T to T_{n-1} in the class of minimal spanning trees. Theorem is proved.

Remark. For a connected weighted graph with n vertices and m edges the complexity of the Kruskal's algorithm can be naturally estimated as $mn \sim n^3$. The estimation can be improved to $m \log m \sim n^2 \log n$. The fastest non-randomized comparison-based algorithm with known complexity belongs to Bernard Chazelle [5]. It turns out that if the weight function is geometrical, then the algorithms can be improved.

2.2 Euclidean Case: Geometrical Approach

Now assume that M is a finite subset of the Euclidean plane \mathbb{R}^2 . It turns out that a minimal spanning tree for M in \mathbb{R}^2 can be constructed faster than the one for an abstract complete graph with $n = |M|$ vertices by means of some geometrical reasonings. To do that we need to construct so called Voronoi partition of the plane, corresponding to M , and the Delaunay graph on M . It turns out that any minimal spanning tree for M in \mathbb{R}^2 is a subgraph of the Delaunay graph, see Figure 3, and the number of edges in this graph is linear with respect to n , so the standard Kruskal's algorithm applied to it gives the complexity $n \log n$ instead of $n^2 \log n$ for the complete graph with n vertices.

Let us pass to details. Let $M = \{A_1, \dots, A_n\} \subset \mathbb{R}^2$ be a finite subset of the plain. The *Voronoi cell* of the point A_i is defined as

$$\text{Vor}_M(A_i) = \{x \in \mathbb{R}^2 \mid \|x - A_i\| \leq \|x - A_j\| \text{ for all } j\}.$$

The Voronoi cell for A_i is a convex polygonal domain which is equal to the intersection of the closed half-planes restricted by the perpendicular bisectors of the segments $A_i A_j$, $j \neq i$. It is easy to verify, that the intersection of any two Voronoi cells has no interior points and that $\cup_i \text{Vor}_M(A_i) = \mathbb{R}^2$. This partition of the plane is referred as *Voronoi partition* or *Voronoi diagram*. Two cells $\text{Vor}_M(A_i)$ and $\text{Vor}_M(A_j)$ are said to be *adjacent*, if there intersection contains a straight segment. The *Delaunay graph* $D(M)$ is defined as the dual planar graph to the Voronoi diagram. More precisely, the vertex set of $D(M)$ is M , and to vertices A_i and A_j are connected by an edge, if and only if their Voronoi cells $\text{Vor}_M(A_i)$ and $\text{Vor}_M(A_j)$ are adjacent. The edges of the Delaunay graph are the corresponding straight segments.

It is easy to verify, that if the set M is generic in the sense that no three points lie at a common straight line and no four points lie at a common circle, then the Delaunay graph $D(M)$ is a triangulation, i.e. its bounded faces are triangles. In general case some bounded faces could be inscribed polygons. Anyway, the number of edges of the graph $D(M)$ does not exceed $3n$. It remains to prove the following key Lemma.

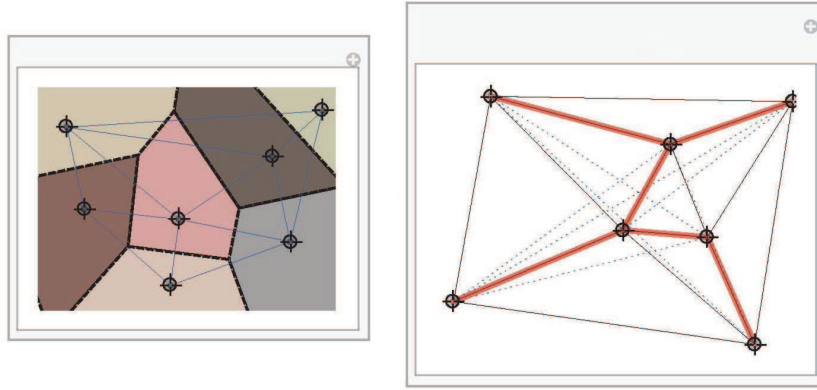


Figure 3: Voronoi diagram (left) and Delaunay graph together with a minimal spanning tree (right) for a point set in the plane.

Lemma 2.4 *Any minimal spanning tree for $M \subset \mathbb{R}^2$ is a subgraph of the Delaunay graph $D(M)$.*

Proof. Let $e = A_i A_j$ be an edge of a minimal spanning tree T for M . We have to show that the Voronoi cells $\text{Vor}_M(A_i)$ and $\text{Vor}_M(A_j)$ are adjacent. The graph $T \setminus e$ consists of two connected components, and this partition generates a partition of the set M into two subsets, say M_1 and M_2 . Assume that $A_i \in M_1$ and $A_j \in M_2$. The minimality of the spanning tree T implies that $\|A_i, A_j\|$ is equal to the distance between the sets M_1 and M_2 , where $\|A_i, A_j\|$ stands for the distance between A_i and A_j .

By u we denote the middle point of the straight segment $A_i A_j$, and let A_k be another point from M . Assume that $A_k \in M_2$. Due to the previous remark, $\|A_i, A_k\| \geq \|A_i, A_j\|$, therefore

$$\|u, A_k\| \geq \|A_i, A_k\| - \|A_i, u\| \geq \|A_i A_j\| - \|A_i, u\| = \|A_i, A_j\|/2 = \|u, A_i\| = \|u, A_j\|.$$

On the other hand, if $\|u, A_k\| = \|u, A_i\|$, then we have equalities in both above inequalities. The first one means that u lies at the straight segment $A_i A_k$, hence A_k lies at the ray $A_i u$. The second equality implies $\|A_i A_k\| = \|A_i A_j\|$, and so $A_k = A_j$, a contradiction. Thus, $\|u, A_k\| > \|u, A_i\|$, that is u does not belong to the cell $\text{Vor}_M(A_k)$ for $k \notin \{i, j\}$. Thus, $u \in \text{Vor}_M(A_i) \cap \text{Vor}_M(A_j)$.

Since the inequality proved is strict, the same arguments remain valid for points lying close to u on the perpendicular bisector to the segment $A_i A_j$. Therefore, the intersection of the Voronoi cells $\text{Vor}_M(A_i)$ and $\text{Vor}_M(A_j)$ contains a straight segment, that is the cells are adjacent. Lemma is proved.

Remark. The previous arguments work in any dimension. But the trouble is that starting from the dimension 3 the number of edges in the Delaunay graph need not be linear on the number of its vertices.

Exercise 2.5 *Verify that the same arguments can be applied to minimal spanning trees for a finite subset of \mathbb{R}^n .*

Exercise 2.6 *Give an example of a finite subset $M \subset \mathbb{R}^3$ such that the Delaunay graph $D(M)$ coincides with the complete graph $K(M)$.*

Problem 2.7 *In what metric spaces similar geometrical approach also works? It definitely works for planar polygons with intrinsic metric, see [6].*

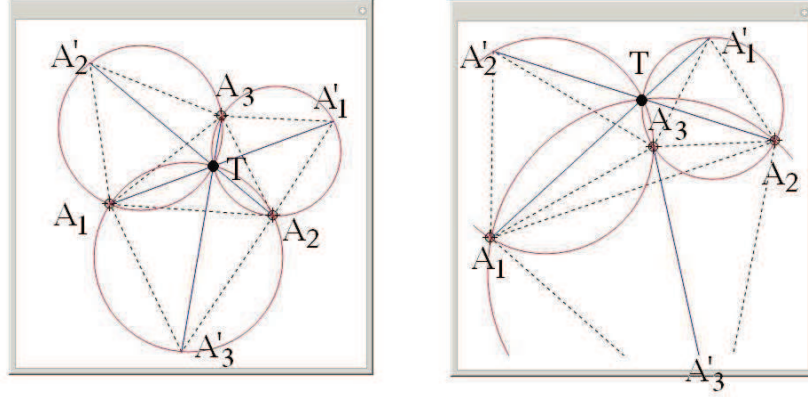


Figure 4: Torricelli–Simpson construction, the case $A_i \leq 120^\circ$ (left), and the case $A_3 > 120^\circ$ (right).

3 Steiner Trees and Locally Minimal Networks

In this section we speak about shortest trees and locally shortest networks in more details. Besides necessary definitions we discuss local structure theorems, Melzak algorithm constructing locally minimal trees in the plane, global results concerning locally minimal binary trees in the plane (so called twisting number theory) and the particular case, locally minimal binary trees with convex boundaries (language of triangular tilings). The details concerning twisting number and tiling realization theory can be found in [2] or [1], and also in [7].

3.1 Fermat Problem

The idea that additional forks can help to decrease the length of a connecting network had been already clear to P. Fermat and his students. It seems that Fermat was the first, who stated the following optimization problem: for given three points A_1 , A_2 , and A_3 in the plane find a point X minimizing the sum of distances from the points A_i , i.e. minimize the function $F(X) = \sum_i \|A_i, X\|$. For the case when all the angles of the triangle $A_1A_2A_3$ are less than or equal to 120° the solution was found by E. Torricelli and later by R. Simpson. The construction of Torricelli is as follows, see Figure 4.

On the sides of the triangle $A_1A_2A_3$ construct equilateral triangles $A_iA_jA'_k$, $\{i, j, k\} = \{1, 2, 3\}$, such that they intersect the initial triangle only by the common sides. Then, as Torricelli proved, the circumscribing circles of these three triangles intersect in a point referred as *Torricelli point T of the triangle $A_1A_2A_3$* . If all the angles A_i are less than or equal to 120° , then T lies in the triangle $A_1A_2A_3$ and gives the unique solution to the Fermat problem.² Later Simpson proved that the straight segments $A_iA'_i$ also pass through the Torricelli point, and the lengths of all these three segments are equal to $F(T)$. If one of the angles, say A_3 , is more than 120° , then the Torricelli point is located outside the triangle and can not be the solution to Fermat problem. In this case the solution is $X = A_3$.

² An elementary proof can be obtained by rotation R of a copy of the triangle around its vertex, say A_1 , by 60° and considering the polygonal line L joining A_2 , X , image $R(X)$ of X under the rotation, and $R(A_3)$. The length of L is equal to $F(X)$, and minimal value of $F(X)$ corresponds to the location of the X such that L is a straight segment.

Remark. So we see, that shortest tree for a triangle in the plane consists of straight segments meeting at the vertices by angles more than or equal to 120° . It turns out, that this 120° -property remains valid in much more general situation.

3.2 Local Structure Theorem and Locally Minimal Networks

Let $M = \{A_1, \dots, A_n\}$ be a finite subset of Euclidean space \mathbb{R}^N , and T is a Steiner tree connecting M . Recall that we defined shortest trees as abstract graphs with vertex set in the ambient metric space. In the case of \mathbb{R}^N it is natural to model edges of such graph as straight segments joining corresponding points in the space. The configuration obtained is referred as a *geometrical realization of the corresponding graph*. Below, speaking about shortest trees in \mathbb{R}^N we will usually mean their geometrical realizations. The local structure of a shortest tree (more exactly of a geometrical realization of the tree) can be easily described.

Theorem 3.1 (Local Structure) *Let Γ be a shortest tree connecting a finite subset $M = \{A_1, \dots, A_n\}$ in \mathbb{R}^N . Then*

1. *all edges of Γ are straight segments;*
2. *any vertex $v \in \Gamma$ of degree 1 belongs to M ;*
3. *any two neighboring edges of Γ meet in common vertex by angle more than or equal to 120° ;*
4. *if the degree of a vertex v is equal to 2 and $v \notin M$, then the edges meet at v by 180° angle.*

Corollary 3.2 *Let Γ be a shortest tree connecting a finite subset $M = \{A_1, \dots, A_n\}$ in \mathbb{R}^N . Then the degree of any its vertex is at most 3, and if the degree of a vertex v equals to 3, then the edges meet at v by angles equal to 120° .*

Example. Let M be the vertex set of regular tetrahedron Δ in \mathbb{R}^3 . Then the network consisting of four straight segments joining the vertices of the tetrahedra with its center O is not a shortest network. Indeed, since $\deg O = 4$, then the angles between the edges meeting at O are less than 120° . The set M is connected by three different (but isometrical) shortest networks, each of which has two additional vertices of degree 3, see Figure 5.

Theorem 3.1 can be just “word-by-word” extended to the case of Riemannian manifolds (we only need to change straight segments by geodesic segments) [1] and even to the case of Alexandrov spaces with bounded curvature. The case of normed spaces turned out to be more complicated (some general results can be found in [2]).

A connected graph Γ in \mathbb{R}^N (in a Riemannian manifold) whose vertex set contains a finite subset $M \subset \mathbb{R}^N$ is called a *locally minimal network connecting M* or *with the boundary $\partial\Gamma = M$* , if it satisfies Conditions (1)–(4) from Theorem 3.1. In the case of complete Riemannian manifolds such graphs are minimal “in small,” i.e. the following result holds, see [1].

Theorem 3.3 (Minimality “in small”) *Let Γ be a locally minimal network connecting a finite subset M of a complete Riemannian manifold W . Then each point $P \in \Gamma$ possesses a neighborhood U in W , such that the network $\Gamma \cap U$ is a shortest network with the boundary $(\partial\Gamma \cap U) \cup (\Gamma \cap \partial U)$.*

Remark. In the case of normed spaces Theorem 3.3 is not valid even for two-point sets, see example in Figure 1.

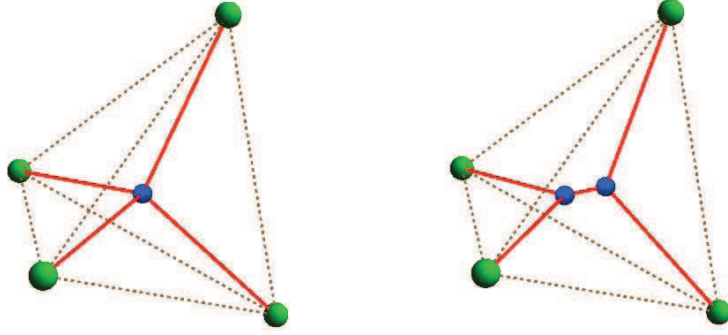


Figure 5: Non-shortest tree (left) and one of the shortest trees (right) for the vertex set of a regular tetrahedron.

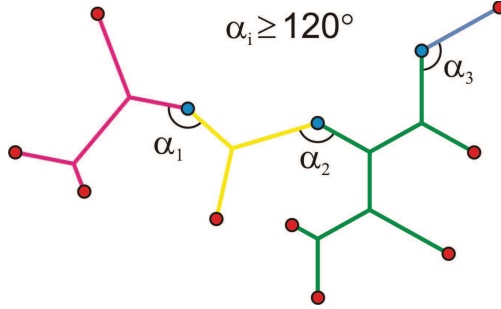


Figure 6: Tree G is partitioned into 4 non-degenerate components by cutting at vertices of degree 2.

3.3 Melzak Algorithm and Steiner Problem Complexity

Let us return back to the case of Euclidean plane. It turns out that in this case the Torricelli–Simpson construction can be generalized to a geometrical algorithm, that either constructs a locally minimal tree of a given structure for a given boundary set, or reports that such a tree does not exist. This algorithm was discovered by Z. Melzak [9] and improved by F. Hwang [8].

Assume that we are given with a tree G whose vertex degrees are at most 3, a finite subset M of the plane, and a bijection $\varphi: \partial G \rightarrow M$, where ∂G is the set of all vertices from G of degrees 1 and 2. To start with, partition the tree G into the union of so-called *non-degenerate* components G_i by cutting the tree at each its vertex of degree 2, see Figure 6. To construct locally minimal network Γ of type G spanning M in accordance with φ it suffices to construct each its component Γ_i of type G_i on the corresponding boundary $M_i = \varphi(\partial G_i)$, where $\partial G_i = \partial G \cap G_i$, in accordance with $\varphi_i = \varphi|_{\partial G_i}$ and to verify the angles between the edges of the components at the vertices of degree 2. All these angles must be more than or equal to 120° , see Figure 6.

Now we pass to the case of one non-degenerate component, i.e. we assume that G has no vertices of degree 2 and that ∂G consists of all the vertices of degree 1. Such trees are referred as *binary*. If $|\partial G| = 2$, then the corresponding locally minimal tree Γ

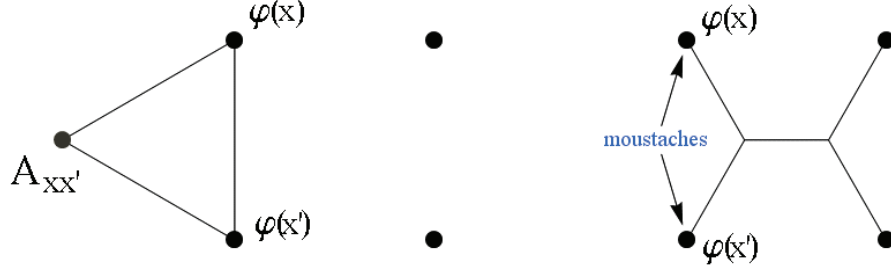


Figure 7: A step of forward trace of Melzak algorithm.

is a straight segment. Otherwise, it is easy to verify that each such tree G contains so-called *moustaches*, i.e. a pair of vertices of degree 1 neighboring with a common vertex of degree 3. Fix such moustaches $\{x, x'\} \subset \partial G$, by y denote their common vertex of degree 3, and make a *forward step of Melzak algorithm*, see Figure 7, that reduces the number of boundary vertices by 1. Namely, we reconstruct the tree G by deleting the vertices x and x' together with the edges xy and $x'y$ and adding y to the boundary of new binary tree; reconstruct the set M by deleting the points $\varphi(x)$ and $\varphi(x')$ and adding a new point $A_{xx'}$ which is the third vertex of a regular triangle constructed on the straight segment $\varphi(x)\varphi(x')$ in the plane; and reconstruct the mapping φ putting $\varphi(y) = A_{xx'}$. Notice that the point $A_{xx'}$ can be constructed in two ways, because there are two such regular triangles. Thus, if the number of boundary vertices in the resulting tree is more than 2, then we can repeat the procedure described above. And if it becomes 2, then we can construct the corresponding locally minimal tree — the straight segment. Here the forward trace of Melzak algorithm stops. Now we have to reconstruct the initial tree, if possible.

Thus, we have a straight segment $I \subset \mathbb{R}^2$ realizing locally minimal tree with unique edge ab , and at least one of its ending points has the form $A_{xx'}$, where x and x' are the boundary vertices of the binary tree G from the previous step, neighboring with their common vertex of degree 3. Let this common vertex be a , that is a corresponds to $A_{xx'}$. We reconstruct G by adding edges ax and ax' . Then we restore the points $\varphi(x)$ and $\varphi(x')$ in the plane together with the regular triangle $\varphi(x)\varphi(x')A_{xx'}$, circumscribe the circle S^1 around it and consider the intersection of S^1 with the segment I , see Figure 8. If it does not contains a point lying at the smaller arc of S^1 restricted by $\varphi(x)$ and $\varphi(x')$, then the tree G can not be reconstructed and we have to pass to another realization of the forward trace of the algorithm. Otherwise we put $\varphi(a)$ be equal to this point. The straight segments $\varphi(x)\varphi(a)$ and $\varphi(x')\varphi(a)$ meet at $\varphi(a)$ by 120° and together with the subsegment $\varphi(a)\varphi(b)$ form a locally minimal binary tree Γ of type G with tree boundary vertices. We repeat this procedure until we either reconstruct the tree of type G , or verify all possible realizations of the forward trace and conclude that the tree of type G does not exists.

The Melzak algorithm described above contains an exponential number of possibilities of its forward trace realization, due to two possible locations of each regular triangle constructed by the algorithm. This complexity can be reduced by means of modification suggested by F. Hwang [8]. He showed that considering a bit more complicated configurations of boundary points (four points corresponding to “neighboring moustaches” or three points corresponding to moustaches and “neighboring” degree-1 vertex) one always can understand which regular triangle must be chosen, see details in [8].

But unfortunately even a linear time realization of Melzak algorithm does not lead to

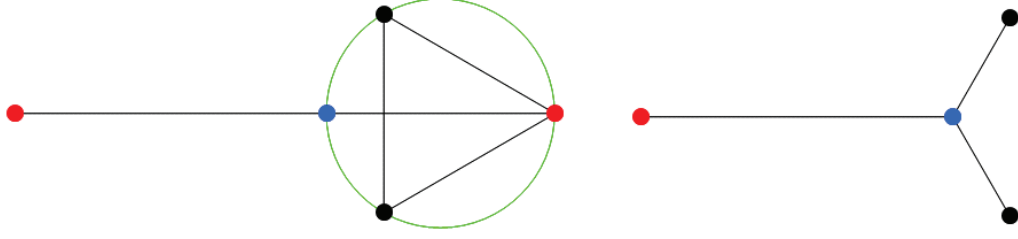


Figure 8: A step of back trace of Melzak algorithm.

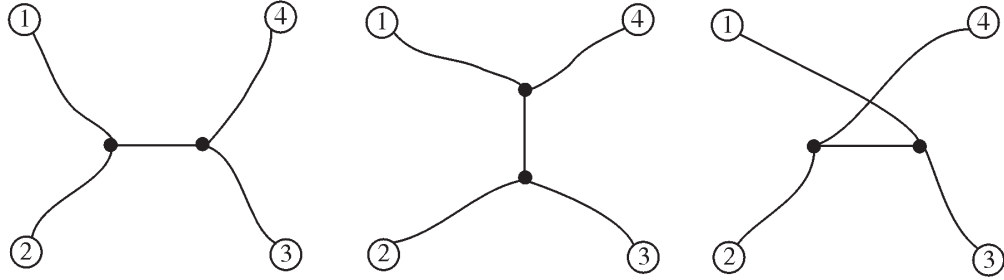


Figure 9: Three binary trees with 4 vertices of degree 1.

a polynomial algorithm of a shortest tree finding. The reason is a huge number of possible structures of the tree G with $|\partial G| = n$ together with also exponential number of different mappings $\varphi: \partial G \rightarrow M$ for fixed ∂G and M . Even for binary trees we have 3 possibilities for $n = 4$, see Figure 9, and 15 possibilities for $n = 5$ (notice that the corresponding binary trees are isomorphic as graphs). For $n = 6$ we have two non-isomorphic binary trees and the number of possibilities becomes 90. It can be shown that the total number of possibilities can be estimated by Catalan number and grows exponentially.

So, to obtain an efficient algorithm, we have to find some *a priori* restrictions on possible structures of minimal networks. In the next subsection we tell about the restrictions generated by geometry of boundary sets.

3.4 Boundaries Geometry and Networks Topology

Here we review our results from [10] and [7]. The goal is to find some restriction on the structure of locally minimal binary trees spanning a given boundary in the plane in terms of geometry of the boundary set. To do this we need to choose or to introduce some characteristics of the network structure and of the boundary geometry.

As a characteristic of the geometry of a boundary set M we take the *number of convexity levels* $c(M)$. Recall the definition. Let M be a finite non-empty subset of the plane. Take the convex hull $\text{ch} M$ of M and assign the points from M lying at the boundary of the polygon $\text{ch} M$ to the *first convexity level* $M^{(1)}$ of M . If the set $M \setminus M^{(1)}$ is not empty, then define the *second convexity level* $M^{(2)}$ of M to be equal to the first convexity level of $M \setminus M^{(1)}$, and so on. As a result, we obtain the partition of the set M into its convexity levels, and by $c(M)$ we denote the total number of this levels, see

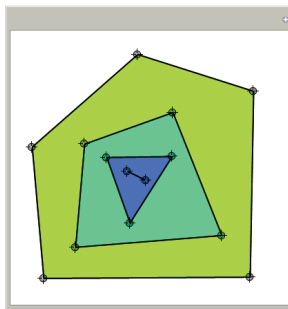


Figure 10: A boundary set with 4 convexity levels.

Figure 10.

Now let us pass to definition of a characteristic describing the “complexity” of planar binary trees. Assume that we are given with a planar binary tree Γ , and let the orientation of the plane be fixed. For any its two edges, say e_1 and e_2 , we consider the unique path γ in Γ starting at e_1 and finishing at e_2 . All interior vertices of γ are the vertices of Γ having degree 3. Let us walk from e_1 to e_2 along γ . Then at each interior vertex of γ we make either left, or right turn in Γ . Define the value $\text{tw}(e_1, e_2)$ to be equal to the difference between the numbers of left and right turns we have made. In other words, assign to an interior vertex of γ the label $\tau = \pm 1$, where $+1$ corresponds to left turns and -1 to right turns. Then $\text{tw}(e_1, e_2)$ is the sum of these values, see Figure 11. Notice that $\text{tw}(e_1, e_2) = -\text{tw}(e_2, e_1)$. At last, we put $\text{tw } \Gamma = \max \text{tw}(e_i, e_j)$, where the maximum is taken over all ordered pairs of edges of Γ .

If the tree Γ is locally minimal, then the twisting number between any pair of its edges has a simple geometrical interpretation, see Figure 11. Namely, since the angles between any neighboring edges are equal to $2\pi/3$, then $\text{tw}(e_i, e_j)$ is equal to the total angle which the oriented edge rotates by passing from e_i to e_j , divided by $\pi/3$.

It turns out, that the twisting number of a locally minimal binary tree with a given boundary is restricted from above by a linear function on the number of convexity levels of the boundary. Namely, the following result holds.

Theorem 3.4 *Let Γ be a locally minimal binary tree connecting the boundary set M that coincides with the set of vertices of degree 1 from Γ . Then*

$$\text{tw } \Gamma \leq 12(c(M) - 1) + 5.$$

The important particular case $c(M) = 1$ corresponds to the vertex sets of convex polygons. Such boundaries are referred as *convex*.

Theorem 3.5 *Let Γ be a locally minimal binary tree with a convex boundary. Then $\text{tw } \Gamma \leq 5$. Conversely, any planar binary tree Γ with $\text{tw } \Gamma \leq 5$ is planar equivalent to a locally minimal binary tree with a convex boundary.*

Notice that the direct statement of Theorem 3.5 is rather easy to prove (it follows from the geometrical interpretation of the twisting number, easy remark that $\text{tw } \Gamma$ always attains at boundary edges, and the monotony of convex polygonal lines). But the converse statement is quite non-trivial. The proof obtained in [7] is based on the complete description of binary trees with twisting number at most five, obtained in terms of so-called triangular tilings that will be discussed in the next subsection.

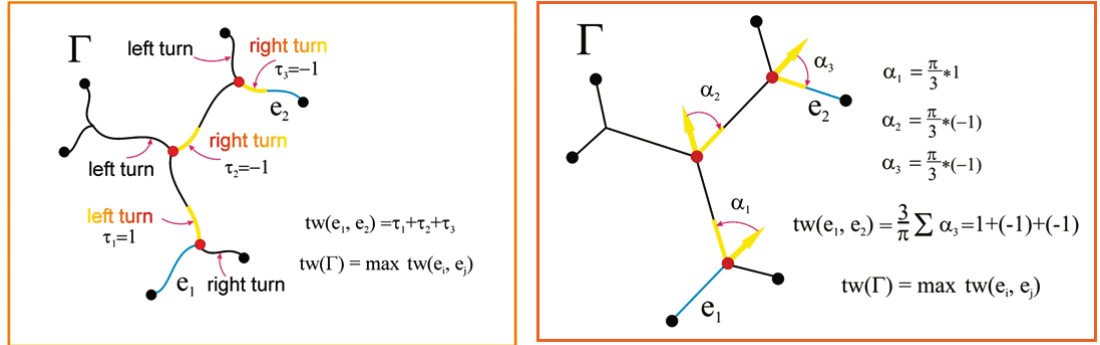


Figure 11: Left and right turns in planar binary tree and its twisting number (left), and twisting number of locally minimal binary tree.

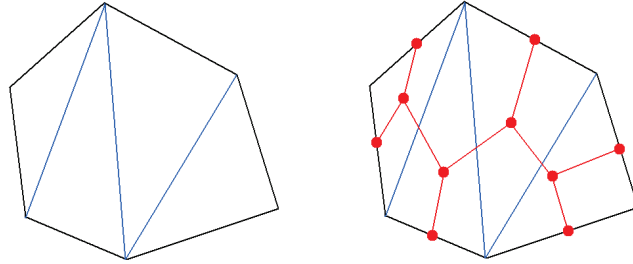


Figure 12: A diagonal triangulation and corresponding planar binary tree.

Problem 3.6 *Estimate the number of binary trees structures with n vertices of degree 1 and twisting number at most k . It is more or less clear that the number is exponential on n even for $k = 5$, but it is interesting to obtain an exact asymptotic.*

3.5 Triangular Tilings and their Applications

It turns out that the description of planar binary trees with twisting number at most five can be effectively done in the language of planar triangulations of a special type which are referred as triangular tilings.

The correspondence between diagonal triangulations of planar convex polygons and planar binary trees is well-known: the planar dual graph of such triangulation is a binary tree, see Figure 12, and each binary tree can be obtained in such a way. Here the vertices of the dual graph are centers of the triangles of the triangulation (medians intersection point) and middle points of the sides of the polygon; and edges are straight segments joining either the middle of a side with the center of the same triangle, or two centers of the triangles having a common side.

In the context of locally minimal binary trees, the most effective way to represent the diagonal triangulations is to draw them consisting of regular triangles. Such special triangulations are referred as *triangular tilings*. The main advantage of the tilings is that the dual binary tree constructed as described above is a locally minimal binary tree with the corresponding boundary. Therefore, tilings “feel the geometry” of locally minimal binary trees and turns out to be very useful in the description of such trees with small

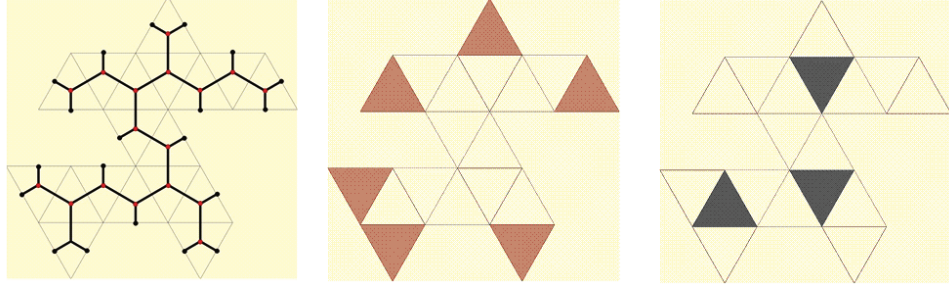


Figure 13: A triangular tiling with its dual binary tree (left), its outer cells (middle) and inner cells (right).

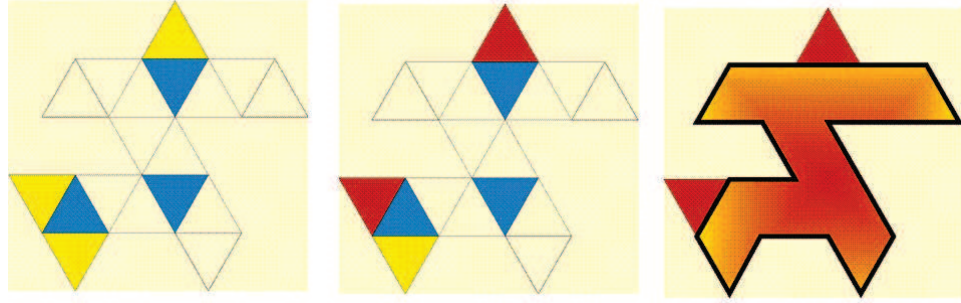


Figure 14: Un-paired and paired growths of a tiling (left), growths that should be deleted to get a skeleton (middle), corresponding decomposition into skeleton and growths (right).

twisting numbers.

The main difficulty in constructing a triangulation consisting of regular triangles for a given binary tree is that the resulting polygon can overlap itself. An example of such overlapping can be easily constructed from a binary tree Γ corresponding to the diagonal triangulation of a convex n -gon, $n \geq 6$, all whose diagonals are incident to a common vertex. But the twisting number of such Γ is also at least 6. The following result is proved in [7].

Theorem 3.7 *The triangular tiling corresponding to any planar binary tree with twisting number less than or equal to five has no self-intersections.*

Theorem 3.7 gives an opportunity to reduce the description of the planar binary trees with twisting number at most five to the description of the corresponding triangular tilings.

To describe all the triangular tilings whose dual binary trees have the twisting number at most five, we decompose each such tiling into elementary “breaks”. The triangles of the tiling are referred as *cells*. A cell of a tiling T is said to be *outer*, if two its sides lie at the boundary of T considered as planar polygon. Further, a cell is said to be *inner*, if no one of its sides lies at the boundary, see Figure 13. An outer cell adjacent to (i.e. intersecting with by a common side) an inner cell is referred as a *growth* of T . A tiling can contain as un-paired growths, so as paired growths, see Figure 14.

For each inner cell we delete exactly one growth adjacent to it, providing such growths

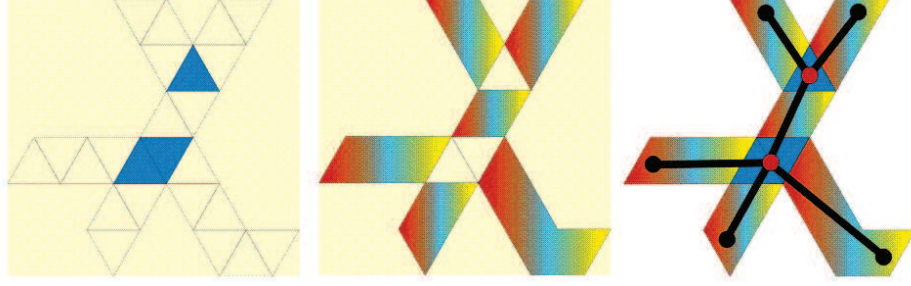


Figure 15: Branching points (left), linear parts (middle), and code of a skeleton (right).

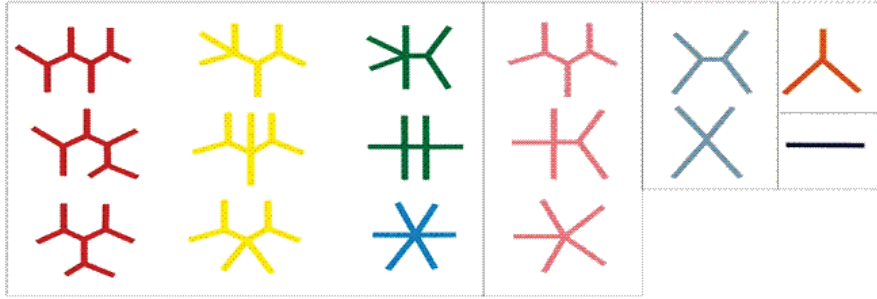


Figure 16: All possible codes of skeletons whose dual binary trees have twisting number at most 5.

exist. As a result, we obtain a decomposition of the initial tiling into its growths and its *skeleton* (a tiling without growths). Notice, that such a decomposition is not unique.

It turns out that the skeletons of the tilings whose dual binary trees have twisting number at most five can be described easily. Also, the possible location of growths in such tilings on their skeletons also can be described. The details can be found in [7] or [1]. Here we only formulate the skeletons describing Theorem and include several examples of its application.

Inner cells of a skeleton S are organized into so-called *branching points*, see Figure 15. After the branching points deleting, the skeleton is partitioned into *linear parts*. Each linear part contains at most one outer cell. Construct a graph $C(S)$ referred as the *code of the skeleton* S as follows: the vertex set of $C(S)$ is the set of its branching points and of the outer cells of its linear parts. The edges correspond to the linear parts, see Figure 15.

The following result is proved in [7].

Theorem 3.8 *Consider all skeletons whose dual graphs twisting numbers are at most 5 and for each of these skeletons construct its code. Then, up to planar equivalence, we obtain all planar graphs with at most 6 vertices of degree 1 and without vertices of degree 2. In particular, every such skeleton contains at most 4 branching points and at most 9 linear parts.*

All possible codes of such skeletons are depicted in Figure 16.

This description of skeletons and corresponding tilings obtained in [7], was applied to the proof of inverse (non-trivial) statement of Theorem 3.5. In some sense, the proof

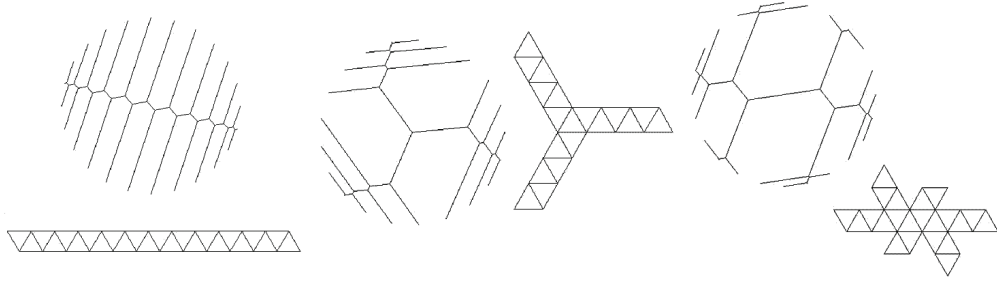


Figure 17: Two infinite families of locally minimal binary trees connecting vertex sets of regular polygons, existing for any regular n -gon (left) and for $3k + 6$ -gons (middle), and the unique finite family, existing for 24-, 30-, 36- and 42-gons only (right).

obtained in [7] is constructive: for each tiling under consideration a corresponding locally minimal binary tree with a convex boundary is constructed.

Another application is a description of all possible binary trees of the skeleton type that can be realized as locally minimal binary trees connecting the vertex set of a regular polygon. It turns out, see details in [1], that there are 2 infinite families of such trees and 1 finite family. The representatives of these networks together with the corresponding skeletons are shown in Figure 17.

4 Steiner Ratio

As we have already discussed in the previous Section, the problem of finding a shortest tree connecting a given boundary set is exponential even in two-dimensional Euclidean plane. On the other hand, in practice it is necessary to solve transportation problems of this kind for several thousands boundary points many times a day. Therefore, in practice some heuristical algorithms are used. One of the most popular heuristics for a shortest tree is corresponding minimal spanning tree. But using such approximate solutions instead of exact one it is important to know the value of possible error appearing under the approximation. The *Steiner ratio of a metric space* is just the measure of maximal possible relative error for the approximation of a shortest tree by the corresponding minimal spanning tree.

4.1 Steiner Ratio of a Metric Space

Let M be a finite subset of a metric space (X, ρ) , and assume that $|M| \geq 2$. We put $\text{sr } M = \text{smt}(M) / \text{mst}(M)$. Evidently, $\text{sr } M \leq 1$. The next statement is also easy to prove.

Assertion 4.1 *For any metric space (X, ρ) and any its finite subset $M \subset X$, $|M| \geq 2$, the inequality $\text{sr } M > 1/2$ is valid.*

Proof. Let G be a Steiner tree connecting M . Consider an arbitrary embedding of the graph G into the plane, walk around G in the plane and list consecutive paths forming this tour and joining consecutive boundary vertices from M . The length of each such path γ_{PQ} joining boundary vertices PQ , i.e. the sum of the lengths of its edges, is more than or equal to the distance $\rho(P, Q)$, due to the triangle inequality. Consider the cyclic path in the complete graph with vertex set M consisting of edges formed by the pairs of consecutive vertices from the tour, and let T be a spanning tree on M contained in

this path. It is clear, that $\rho(T) < \sum_{(P,Q)} \rho(\gamma_{PQ})$, where the summation is taken over all the pairs of consecutive vertices of the tour. On the other hand, each edge of the tree G belongs to exactly two such paths, hence $\sum_{(P,Q)} \rho(\gamma_{PQ}) = 2\rho(G)$. So, we have $\text{sr}(M) \geq \rho(G)/\rho(T) > 1/2$. The Assertion is proved.

The value $\text{sr}(M)$ is the relative error appearing under approximation of the length of a shortest tree for a given set M by the length of a minimal spanning tree. The *Steiner ratio of a metric space* (X, ρ) is defined as the value $\text{sr}(X) = \inf_{M \subset X} \text{sr}(M)$, where the infimum is taken over all finite subsets M , $|M| \geq 2$ of the metric space X . So, the Steiner ratio of X is the value of the relative error in the worse possible case.

Corollary 4.2 *For arbitrary metric space (X, ρ) the inequality $1/2 \leq \text{sr}(X) \leq 1$ is valid.*

Exercise 4.3 *Verify, that for any $r \in [1/2, 1]$ there exists a metric space (X, ρ) with $\text{sr}(X) = r$, see corresponding examples in [2].*

Sometimes, it is convenient to consider so-called *Steiner ratios* $\text{sr}_n(X)$ of degree n , where $n \geq 2$ is an integer, which are defined as follows: $\text{sr}_n(X) = \inf_{M \subset X, |M| \leq n} \text{sr}(M)$. Evidently, $\text{sr}_2(X) = 1$. It is also clear that $\text{sr}(X) = \inf_n \text{sr}_n(X)$.

Steiner ratio was firstly defined for the Euclidean plane in [11], and during the following years the problem of Steiner ratio calculation is one of the most attractive, interesting and difficult problems in geometrical optimization. A short review can be found in [2] and in [12]. One of the most famous stories here is connected with several attempts to prove so-called *Gilbert–Pollack Conjecture*, see [11], saying that $\text{sr}(\mathbb{R}^2, \rho_2) = \sqrt{3}/2$, where ρ_2 stands for the Euclidean metric, and hence $\text{sr}(\mathbb{R}^2, \rho_2)$ is attained at the vertex set of a regular triangle, see Figure 2. In 1990s D. Z. Du and F. K. Hwang announced that they proved the Steiner Ratio Gilbert–Pollak Conjecture [13], and their proof was published in *Algorithmica* [14]. In spite of the appealing ideas of the paper, the questions concerning the proof appeared just after the publication, because the text did not appear formal. And about 2003–2005 it becomes clear that the gaps in the D. Z. Du and F. K. Hwang work are too deep and can not be repaired, see detail in [15].

4.2 Steiner Ratio of Small Degrees for Euclidean Plane

Gilbert and Pollack calculated $\text{sr}_3(\mathbb{R}^2, \rho_2)$ in their paper [11]. We include their proof here.

Since the Steiner ratio of a regular triangle is equal to $\sqrt{3}/2$, then $\text{sr}_3(\mathbb{R}^2, \rho_2) \leq \sqrt{3}/2$, so we just need to prove the opposite inequality. To do this, consider a triangle ABC in the plane. If one of its angles is more than or equal to 120° , then the shortest tree coincides with minimal spanning tree, so in this case $\text{sr}(ABC) = 1$. So it suffices to consider the case when all the angles of the triangle are less than 120° .

Let S be the Torricelli point of the triangle ABC . Show firstly that $|AS| \leq |BS|$, if and only if $|BC| \geq |AC|$, i.e. the shortest edge of the Steiner minimal tree lies opposite with the longest side of the triangle. The proof is shown in Figure 18, left. Indeed, if $|BS| < |AS|$, then we take the point $B' \in [S, B]$ with $|SB'| = |SA|$, hence $|CB'| = |CA|$ due to symmetry and $|CB'| < |CB|$ because $B' \geq 120^\circ$. Conversely, if $|BC| > |B'C|$, then there exists $B' \in [B, S]$ with $|CB'| = |CA|$, because $|BC| > |CA| > |SC|$. Then $|AS| = |SB'| < |SB|$.

Thus, the two-edges tree $T = [A, B] \cup [B, C]$ is a minimal spanning tree for ABC , if and only if BC is the longest side of ABC , if and only if $|AS| \leq |BS|$ and $|AS| \leq |CS|$. Consider the points $E \in [B, S]$ and $D \in [C, S]$, such that $|AS| = |ES| = |DS|$, and put $x = |AC|$, $y = |AB|$, $z = |DE| = |AD| = |AE|$, and $x' = |CD|$, $y' = |EB|$. Then

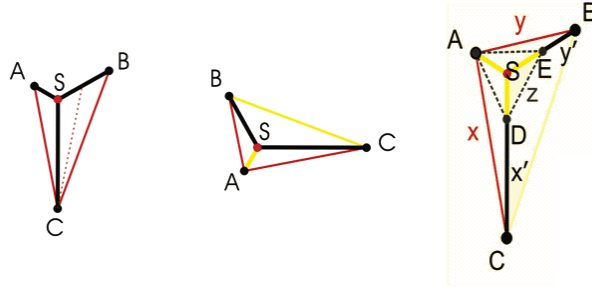


Figure 18: To the calculation of $\text{sr}_3(\mathbb{R}^2, \rho_2)$.

$$|SA| = |SE| = |SD| = z/\sqrt{3} \text{ and}$$

$$\text{smt}(M) = 3|SA| + |DC| + |EB| = \sqrt{3}z + x' + y' \quad \text{and} \quad \text{mst}(M) = x + y,$$

where M stands for the set $\{A, B, C\}$. But $x \leq x' + z$ and $y \leq y' + z$, due to the triangle inequality, and hence

$$\text{sr}(M) = \frac{\sqrt{3}z + x' + y'}{x + y} \geq \frac{\sqrt{3}z + x' + y'}{x' + z + y' + z} = \frac{\sqrt{3}z + x' + y'}{x' + y' + 2z} \geq \frac{\sqrt{3}}{2}.$$

Thus, we proved the following statement.

Assertion 4.4 *The following relation is valid: $\text{sr}_3(\mathbb{R}^2, \rho_2) = \sqrt{3}/2$.*

Remark. For small n it is already proved that $\text{sr}_n(\mathbb{R}^2, \rho_2) = \sqrt{3}/2$ (recently O. de Wet proved it for $n \leq 7$, see [16]). The proof of de Wet is based on the analysis of Du and Hwang method from [14] and understanding that it works for boundary sets with $n \leq 7$ points. Also in 60th several lower bounds for $\text{sr}(\mathbb{R}^2, \rho_2)$ were obtained, and the best of them is worse than $\sqrt{3}/2$ in the third digit only.

Problem 4.5 *Very attractive problem is to prove that $\text{sr}(\mathbb{R}^2, \rho_2) = \sqrt{3}/2$, i.e. to prove Gilbert–Pollack Conjecture. The attempts to repair the proof of Du and Hwang have remained unsuccessful, so some fresh ideas are necessary here.*

4.3 Steiner Ratio of Other Euclidean Spaces and Riemannian Manifolds

The following result is evident, but useful.

Assertion 4.6 *If Y is a subspace of a metric space X , i.e. the distance function on Y is the restriction of the distance function of X , then $\text{sr}(Y) \geq \text{sr}(X)$.*

This implies, that $\text{sr}(\mathbb{R}^n, \rho_2) \leq \text{sr}(\mathbb{R}^2, \rho_2) \leq \sqrt{3}/2$. Recall that Gilbert–Pollack conjecture implies that the Steiner ratio of Euclidean plane attains at the vertex set of a regular triangle. In multidimensional case the situation is more complicated. The following result was obtained by Du and Smith [17]

Assertion 4.7 *If $M \subset \mathbb{R}^n$ is the vertex set of a regular n -dimensional simplex, then $\text{sr}(M) > \text{sr}(\mathbb{R}^n, \rho_2)$ for $n \geq 3$.*

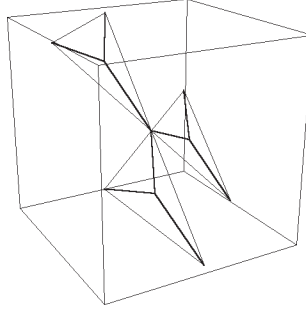


Figure 19: Construction of the set P in \mathbb{R}^3 (non-interesting but visual case $n = 2$).

Proof. Consider the boundary set P in \mathbb{R}^{n+1} , consisting of the following $1 + n(n+1)$ points: one point $(0, \dots, 0)$ and $n(n+1)$ points all whose coordinates except two are zero, one is equal to 1, and the remaining one is -1 . It is clear that P is a subset of n -dimensional plane defined by the next linear condition: sum of all coordinates is equal to zero. Represent P as the union of the subsets $P^i = \{x \in P \mid x^i = 1\} \cup \{(0, \dots, 0)\}$. Notice that each set P^i , $i = 1, \dots, n+1$, consists of $n+1$ points and forms the vertex set of an regular n -dimensional simplex (to see that it suffices to verify that all the distances between the pairs of points from P^i are the same and are equal to $\sqrt{2}$). The configuration of 7 points in \mathbb{R}^3 is shown in Figure 19 (this case is not important for us, but it is easy to draw). Now, $\text{mst}(P) = (n+1) \text{mst}(P^i)$, but for $n \geq 3$ we conclude that $\text{smt}(P) < (n+1) \text{smt}(P^i)$, because the degree of the vertex $(0, \dots, 0)$ in the corresponding network which is the union of the shortest networks for P^i is equal to $n+1 \geq 4$ that is impossible in the shortest network due to the Local Structure Theorem 3.1. So,

$$\text{sr}(P) = \text{smt}(P) / \text{mst}(P) < \frac{(n+1) \text{smt}(P^i)}{(n+1) \text{mst}(P^i)} = \text{sr}(P^i).$$

Taking as a heuristic for the length of a shortest tree connecting the vertex set of regular simplex the length of the network joining the center of the simplex with all its vertices we get the following estimate.

Corollary 4.8 *For any $n \geq 3$ the upper estimate*

$$\text{sr}(\mathbb{R}^n, \rho_2) < \sqrt{\frac{1}{2} + \frac{1}{2n}}$$

is valid.

One of the best general low estimates is obtained by Graham and Hwang in [19].

Assertion 4.9 *For any $n \geq 2$ the lower estimate $1/\sqrt{3} \leq \text{sr}(\mathbb{R}^n, \rho_2)$ is valid.*

The best known upper estimate for \mathbb{R}^3 is obtained by Smith and Smith [18]. It is attained at an infinite boundary set which is known as “Smith sausage” and depicted in Figure 20. The corresponding value, obtained as the limit of the ratios for finite fragments, is as follows:

$$\sqrt{\frac{283}{700} - \frac{3\sqrt{21}}{700} + \frac{9\sqrt{22 - 2\sqrt{21}}}{140}}.$$

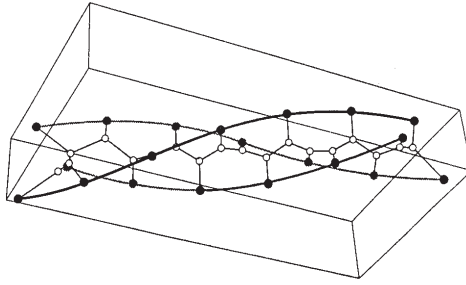


Figure 20: A finite fragment of infinite “Smith sausage”.

Notice that the idea of an infinite set is based on a deep result of Du and Smith estimating from below the number of points in a subset M of \mathbb{R}^n such that $\text{sr}(M) = \text{sr}(\mathbb{R}^n, \rho_2)$ by a function $f(n)$ rapidly increasing on n , see details in [17]. For example, $f(50) = 53$, but $f(200) = 3\,481\,911$. Therefore, it is difficult to expect to guess a finite set M in \mathbb{R}^n with $\text{sr}(M) = \text{sr}(\mathbb{R}^n, \rho_2)$ for large n .

Recently, the Steiner ratio of the Lobachevskii plane, and hence, of any Lobachevskii space has been calculated by Innami and Kim, see [20].

Theorem 4.10 *Steiner ratio of Lobachevskii space L^n for any $n \geq 2$ is equal to $1/2$.*

For general Riemannian manifold Ivanov, Cieslik and Tuzhilin, see [21], obtained the following general result.

Theorem 4.11 *The Steiner ratio of n -dimensional Riemannian manifold is less than or equal to the Steiner ratio of the Euclidean space \mathbb{R}^n .*

5 Minimal Fillings

This Section is devoted to minimal fillings, the third kind of optimal connections discussed in the Introduction. This problem appeared as a result of a synthesis of two classical problems: the Steiner problem on the shortest networks (it is discussed in Sections 3 and 4), and Gromov’s problem on minimal fillings.

The concept of a minimal filling appeared in papers of Gromov, see [22]. Let M be a manifold endowed with a distance function ρ . Consider all possible fillings W spanning M , i.e., compact manifolds with the boundary M . Consider on W a distance function d that does not decrease the distances between points in M . Such a metric space $\mathcal{W} = (W, d)$ is called a *filling* of the metric space $\mathcal{M} = (M, \rho)$, see example in Figure 21. The Gromov Problem consists in calculating the infimum of the volumes of the fillings and describing the spaces \mathcal{W} which this infimum is achieved at (such spaces are called *minimal fillings*).

In the scope of Steiner problem, it is natural to consider M as a finite metric space. Then the possible fillings are metric spaces having the structure of one-dimensional stratified manifolds which can be considered as graphs whose edges have nonnegative weights. This leads to the following particular case of generalized Gromov problem.

Let M be an arbitrary finite set, and $G = (V, E)$ be a connected graph. We say, that G *connects* M or *joins* M , if $M \subset V$. Now, let $\mathcal{M} = (M, \rho)$ be a finite metric space, $G = (V, E)$ be a connected graph joining M , and $\omega: E \rightarrow \mathbb{R}_+$ is a mapping into non-negative numbers, which is usually referred as a *weight function* and which generates the *weighted graph* $\mathcal{G} = (G, \omega)$. The function ω generates on V the pseudo-metric d_ω (some

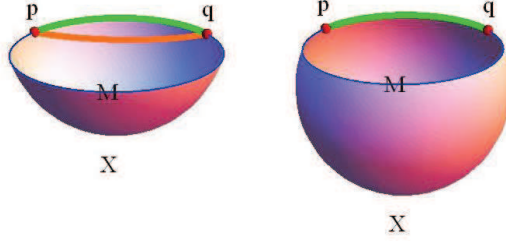


Figure 21: The space M is the circle S^1 with arc-metric. The films X in the both Figures are parts of the standard sphere containing M as a parallel. The left film X is not a filling since the distance between the points p and q in X is less than in M (the shortest path is depicted). The right film X is a filling of M .

distances in a pseudo-metric can be equal to zero), namely, the d_ω -distance between the vertices of the graph \mathcal{G} is defined as the least possible weight of the paths in \mathcal{G} joining these vertices. If for any two points p and q from M the inequality $\rho(p, q) \leq d_\omega(p, q)$ holds, then the weighted graph \mathcal{G} is called a *filling* of the space \mathcal{M} , and the graph G is referred as the *type* of this filling. The value $\text{mf}(\mathcal{M}) = \inf \omega(\mathcal{G})$, where the infimum is taken over all the fillings \mathcal{G} of the space \mathcal{M} is the *weight of minimal filling*, and each filling \mathcal{G} such that $\omega(\mathcal{G}) = \text{mf}(\mathcal{M})$ is called a *minimal filling*.

5.1 Parametric Networks and Optimal Connection Problems

Here we give a common view on Steiner problem and minimal filling problem in terms of so-called parametric networks in a general metric space.

Let $\mathcal{X} = (X, d)$ be a metric space and $G = (V, E)$ be an arbitrary connected graph. Any mapping $\Gamma: V \rightarrow X$ is called a *network in \mathcal{X} parameterized by the graph $G = (V, E)$* , or a *network of the type G* . The *vertices* and *edges* of the network Γ are the restrictions of the mapping Γ onto the vertices and edges of the graph G , respectively. The *length of the edge $\Gamma: vw \rightarrow X$* is the value $d(\Gamma(v), \Gamma(w))$, and the *length $d(\Gamma)$ of the network Γ* is the sum of lengths of all its edges. We shall consider various boundary value problems for graphs. To do that, we fix some subsets ∂G of the vertex sets V of our graphs $G = (V, E)$, and we call such ∂G the *boundaries*. We always suppose that in each graph under consideration a boundary, possibly, an empty one, is chosen. The *boundary $\partial\Gamma$ of a network Γ* is the restriction of Γ onto ∂G . If $M \subset X$ is finite and $M \subset \Gamma(V)$, then we say that the network Γ *joins* or *connects the set M* . The vertices of graphs and networks which are not boundary ones are called *interior* vertices. The value

$$\text{smt}(M) = \inf \{d(\Gamma) \mid \Gamma \text{ is a network joining } M\}$$

is called the *length of shortest network for M* . Notice that the network Γ which joins M and satisfies $d(\Gamma) = \text{smt}(M)$ may not exist, see [23] and [24] for nontrivial examples. If such a network exists, it is called a *shortest network connecting M* , or *for M* . One variant of the Steiner problem is to describe the shortest networks for finite subsets of metric spaces.³

Now let us define minimal parametric networks in a metric space $\mathcal{X} = (X, d)$. Let $G = (V, E)$ be a connected graph with some boundary ∂G , and let $\varphi: \partial G \rightarrow X$ be a

³The denotation *smt* is an acronym for “Steiner Minimal Tree” which is a synonym for the shortest network whose edges are non-degenerate and, thus, it must be a tree.

mapping. By $[G, \varphi]$ we denote the set of all networks $\Gamma: V \rightarrow X$ of the type G such that $\partial\Gamma = \varphi$. We put

$$\text{mpn}(G, \varphi) = \inf_{\Gamma \in [G, \varphi]} d(\Gamma)$$

and we call this value the *length of minimal parametric network*. If there exists a network $\Gamma \in [G, \varphi]$ such that $d(\Gamma) = \text{mpn}(G, \varphi)$, then Γ is called a *minimal parametric network of the type G with the boundary φ* .

Assertion 5.1 *Let $\mathcal{X} = (X, d)$ be an arbitrary metric space and M be a finite subset of X . Then*

$$\text{smt}(M) = \inf \{ \text{mpn}(G, \varphi) \mid \varphi(\partial G) = M \},$$

where the infimum is taken over all connected graphs G with a boundary ∂G and all mappings $\varphi: \partial G \rightarrow X$ with $\varphi(\partial G) = M$.

Thus, as in the case of the plane, the problem of calculating the length of the shortest network is reduced to investigation of minimal parametric networks.

Let $\mathcal{M} = (M, \rho)$ be a finite metric space and $G = (V, E)$ be an arbitrary connected graph connecting M . In this case we always assume that the boundary of such G is fixed and equal to M . By $\Omega(\mathcal{M}, G)$ we denote the set of all weight functions $\omega: E \rightarrow \mathbb{R}$ such that (G, ω) is a filling of the space \mathcal{M} . We put

$$\text{mpf}(\mathcal{M}, G) = \inf_{\omega \in \Omega(\mathcal{M}, G)} \omega(G)$$

and we call this value the *weight of minimal parametric filling of the type G for the space \mathcal{M}* . If there exists a weight function $\omega \in \Omega(\mathcal{M}, G)$ such that $\omega(G) = \text{mpf}(\mathcal{M}, G)$, then (G, ω) is called a *minimal parametric filling of the type G for the space \mathcal{M}* .

Assertion 5.2 *Let $\mathcal{M} = (M, \rho)$ be a finite metric space. Then*

$$\text{mf}(\mathcal{M}) = \inf \{ \text{mpf}(\mathcal{M}, G) \},$$

where the infimum is taken over all connected graphs G joining M .

It is not difficult to show that to investigate shortest networks and minimal fillings one can restrict the consideration to trees such that all their vertices of degree 1 and 2 belong to their boundaries. **In what follows, we always assume that this condition holds, providing the opposite is not declared.**

To be more precise, we recall the following definition. We say that a tree is a *binary* one if the degrees of its vertices can be 1 or 3 only, and the boundary consists just of all the vertices of degree 1. Then each finite metric space has a binary minimal filling (possibly, with some degenerate edges), and a non-degenerate minimal filling (whose type is a tree and all whose vertices of degree 1 and 2 belong to its boundary in accordance with the above agreement), see [3].

5.2 Minimal Realization

It turns out that the problem on minimal filling can be reduced to Steiner problem in special metric spaces and for special boundaries.

Consider a finite set $M = \{p_1, \dots, p_n\}$, and let $\mathcal{M} = (M, \rho)$ be a metric space. We put $\rho_{ij} = \rho(p_i, p_j)$. By \mathbb{R}_∞^n we denote the n -dimensional arithmetic space with the norm

$$\|(v^1, \dots, v^n)\|_\infty = \max\{|v^1|, \dots, |v^n|\},$$

and by ρ_∞ the metric on \mathbb{R}_∞^n generated by $\|\cdot\|_\infty$, i.e., $\rho_\infty(v, w) = \|w - v\|_\infty$. Let us define a mapping $\varphi_{\mathcal{M}}: M \rightarrow \mathbb{R}_\infty^n$ as follows:

$$\varphi_{\mathcal{M}}(p_i) = \bar{p}_i = (\rho_{i1}, \dots, \rho_{in}).$$

Assertion 5.3 *The mapping $\varphi_{\mathcal{M}}$ is an isometry with its image.*

Proof. This easily follows from the triangle inequality. Indeed,

$$\|\bar{p}_i - \bar{p}_j\| = \max_k |\rho_{ik} - \rho_{jk}| \geq \rho_{ij},$$

because the value ρ_{ij} stands at the i th and j th places of the vector $\bar{p}_i - \bar{p}_j$. On the other hand, $\rho_{ij} \geq \rho_{ik} - \rho_{jk}$ for any k , due to the triangle inequality, hence $\|\bar{p}_i - \bar{p}_j\| \leq \rho_{ij}$, and Assertion is proved.

The mapping $\varphi_{\mathcal{M}}$ is called the *Kuratowski isometry*.

Let $\mathcal{G} = (G, \omega)$ be a filling of a space $\mathcal{M} = (M, \rho)$, where $G = (V, E)$, and d_ω be the pseudo-metric on V generated by the weight function ω . Denote by E_M the edges set of the complete graph on M and put $\bar{G} = (V, \bar{E} = E \cup E_M)$. Let $\bar{\omega}$ be the weight function on \bar{E} coinciding with metric ρ on E_M and with ω on $\bar{E} \setminus E_M$. Recall that $d_{\bar{\omega}}$ denotes the pseudo-metric on V generated by $\bar{\omega}$.

We define the network $\Gamma_{\mathcal{G}}: V \rightarrow \mathbb{R}_{\infty}^n$ of the type G as follows:

$$\Gamma_{\mathcal{G}}(v) = (d_{\bar{\omega}}(v, p_1), \dots, d_{\bar{\omega}}(v, p_n)).$$

This network is called the *Kuratowski network for the filling \mathcal{G}* .

Assertion 5.4 *We have $\partial\Gamma_{\mathcal{G}} = \varphi_{\mathcal{M}}$.*

Proof. This easily follows from the filling definition. Indeed, the mapping $\partial\Gamma_{\mathcal{G}}$ is defined on the set M only. By definition,

$$\Gamma_{\mathcal{G}}(p_i) = (d_{\bar{\omega}}(p_i, p_1), \dots, d_{\bar{\omega}}(p_i, p_n)),$$

hence it suffices to show that $d_{\bar{\omega}}(p_i, p_k) = \rho_{ik}$ for any k . The vertices p_i and p_k are joined by the edge $p_i p_k$ of the weight ρ_{ik} in the graph G , and the weight of any other path in G connecting p_i and p_k is more than or equal to ρ_{ik} , because G is a filling. Assertion is proved.

For any network Γ in a metric space (X, d) by ω_{Γ} we denote the *weight function on G induced by the network Γ* , i.e., $\omega_{\Gamma}(vw) = d(\Gamma(v), \Gamma(w))$.

Corollary 5.5 *Let $\mathcal{G} = (G, \omega)$ be a minimal parametric filling of a metric space (M, ρ) and $\Gamma = \Gamma_{\mathcal{G}}$ be the corresponding Kuratowski network. Then $\omega = \omega_{\Gamma}$.*

Let Γ be a network in a metric space \mathcal{X} , let G be its parameterizing graph, and $\mathcal{H} = (H, \omega)$ be a weighted graph. We say that Γ and \mathcal{H} are *isometric*, if there exists an isomorphism of the weighted graphs \mathcal{H} and $\mathcal{G} = (G, \omega_{\Gamma})$.

Corollary 5.5 and the existence of minimal parametric and shortest networks in a finite-dimensional normed space [1] imply the following result.

Corollary 5.6 *Let $\mathcal{M} = (M, \rho)$ be a metric space consisting of n points, and $\varphi_{\mathcal{M}}: M \rightarrow \mathbb{R}_{\infty}^n$ be the Kuratowski isometry. For any graph G joining M there exists a minimal parametric filling of the type G of the space \mathcal{M} . Each minimal parametric filling of the type G of the space \mathcal{M} is isometric to the corresponding Kuratowski network, which is, in this case, a minimal parametric network of the type G with the boundary $\varphi_{\mathcal{M}}$. Conversely, each minimal parametric network of the type G on $\varphi_{\mathcal{M}}(M)$ is isometric to some minimal parametric filling of the type G of the space \mathcal{M} .*

Corollary 5.7 *Let $\mathcal{M} = (M, \rho)$ be a metric space consisting of n points, and $\varphi_{\mathcal{M}}: M \rightarrow \mathbb{R}_{\infty}^n$ be the Kuratowski isometry. Then there exists a minimal filling \mathcal{G} for \mathcal{M} , and the corresponding Kuratowski network $\Gamma_{\mathcal{G}}$ is a shortest network in the space \mathbb{R}_{∞}^n joining the set $\varphi_{\mathcal{M}}(M)$. Conversely, each shortest network on $\varphi_{\mathcal{M}}(M)$ is isometric to some minimal filling of the space \mathcal{M} .*

5.3 Minimal Parametric Fillings and Linear Programming

Let $\mathcal{M} = (M, \rho)$ be a finite metric space connected by a (connected) graph $G = (V, E)$. As above, by $\Omega(\mathcal{M}, G)$ we denote the set consisting of all the weight functions $\omega: E \rightarrow \mathbb{R}_+$ such that $\mathcal{G} = (G, \omega)$ is a filling of \mathcal{M} , and by $\Omega_m(\mathcal{M}, G)$ we denote its subset consisting of the weight functions such that \mathcal{G} is a minimal parametric filling of \mathcal{M} .

Assertion 5.8 *The set $\Omega(\mathcal{M}, G)$ is closed and convex in the linear space \mathbb{R}^E of all the functions on E , and $\Omega_m(\mathcal{M}, G) \subset \Omega(\mathcal{M}, G)$ is a nonempty convex compact.*

Proof. It is easy to see, that the set $\Omega(\mathcal{M}, G) \subset \mathbb{R}^E$ is determined by the linear inequalities of two types: $\omega(e) \geq 0$, $e \in E$, and $\sum_{e \in \gamma_{pq}} \omega(e) \geq \rho(p, q)$, where γ_{pq} stands for the unique path in the tree G connecting the boundary vertices p and q . Therefore, $\Omega(\mathcal{M}, G)$ is a convex closed polyhedral subset of \mathbb{R}^E that is equal to the intersection of the corresponding closed half-spaces. The weight functions of minimal parametric fillings correspond to minima points of the linear function $\sum_{e \in E} \omega(e)$ restricted to the set $\Omega(\mathcal{M}, G)$. Thus, the problem of minimal parametric filling finding is a linear programming problem, and the set $\Omega_m(\mathcal{M}, G)$ of all minima points is a nonempty convex compact polyhedron (the boundedness and, hence, compactness of this set follows from increasing of the objective function with respect to each its variable).

5.4 Generalized Fillings

Investigating the fillings of metric spaces, it turns out to be convenient to expand the class of weighted trees under consideration permitting arbitrary weights of the edges (not only non-negative). The corresponding objects are called *generalized fillings*, *minimal generalized fillings* and *minimal parametric generalized fillings*. Their weights for a metric space \mathcal{M} and a tree G are denoted by $\text{mf}_-(\mathcal{M})$ and $\text{mpf}_-(\mathcal{M}, G)$, respectively.

For any finite metric space $\mathcal{M} = (M, \rho)$ and a tree G connecting M , the next evident inequality is valid: $\text{mpf}_-(\mathcal{M}, G) \leq \text{mpf}(\mathcal{M}, G)$. And it is not difficult to construct an example, when this inequality becomes strict, see Figure 22. However, for minimal generalized fillings the following result holds, see [25].

Theorem 5.9 (Ivanov, Ovsyannikov, Strelkova, Tuzhilin) *For an arbitrary finite metric space \mathcal{M} , the set of all its minimal generalized fillings contains its minimal filling, i.e. a generalized minimal filling with nonnegative weight function. Hence, $\text{mf}_-(\mathcal{M}) = \text{mf}(\mathcal{M})$.*

5.5 Formula for the Weight of Minimal Filling

Let $\mathcal{M} = (M, \rho)$ be a finite metric space, and G be a tree connecting M . Choose an arbitrary embedding G' of the tree G into the plane. Consider a walk around the tree G' . We draw the points of M consecutive with respect to this walk as a consecutive points of the circle S^1 . Notice that each vertex p from M appears $\deg p$ times. For each vertex $p \in M$ of degree more than 1, we choose just one arbitrary point from the corresponding points of the circle. So, we construct an injection $\nu: M \rightarrow S^1$. Define a cyclic permutation π as follows: $\pi(p) = q$, where $\nu(q)$ follows after $\nu(p)$ on the circle S^1 . We say that π is generated by the embedding G' (this procedure is not unique due to different possible choices of ν). Each π generated in this manner is called a *tour of M with respect to G* . The set of all tours on M with respect to G is denoted by $\mathcal{O}(M, G)$. For each tour $\pi \in \mathcal{O}(M, G)$ we put

$$p(\mathcal{M}, G, \pi) = \frac{1}{2} \sum_{x \in M} \rho(x, \pi(x))$$

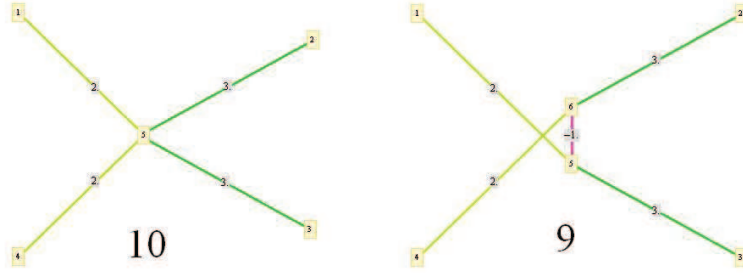


Figure 22: Minimal parametric filling (left) and minimal generalized parametric filling (right) of the vertex set of the plane rectangle with sides 3 and 4. The type is the same: the moustaches connects the diagonal pairs of the vertices. The interior edge has to be zero in the case of the filling and can be negative in the case of the generalized filling. Here $9 = \text{mpf}_-(\mathcal{M}, G) < \text{mpf}(\mathcal{M}, G) = 10$.

and we call this value by the *half-perimeter of the space \mathcal{M} with respect to the tour π* . The minimal value of $p(\mathcal{M}, G, \pi)$ over all $\pi \in \mathcal{O}(M, G)$ for all possible G (in fact, over all possible cyclic permutations π on M) is called the *half-perimeter of the space \mathcal{M}* .

A. Ivanov and A. Tuzhilin proposed the following hypothesis.

Conjecture 5.10 *For an arbitrary metric space $\mathcal{M} = (M, \rho)$ the following formula is valid*

$$\text{mf}(\mathcal{M}) = \min_G \max_{\pi \in \mathcal{O}(M, G)} p(\mathcal{M}, G, \pi),$$

where minimum is taken over all binary trees G connecting M .

A. Yu. Eremin [26] constructed a counter-example to the Conjecture 5.10 and showed that if one changes the concept of tour by the one of multitour, introduced by him, then the Conjecture 5.10 holds.

To define the multitours, let us consider the graph in which every edge of G is taken with the multiplicity $2k$, $k \geq 1$. The resulting graph possesses an Euler cycle consisting of *irreducible* boundary paths — the ones which do not contain properly other boundary paths. This Euler cycle generates a bijection $\pi: X \rightarrow X$, where $X = \sqcup_{i=1}^k M$, which is called *multitour of M with respect to G* , see an example in Figure 23. The set of all multitours on M with respect to G is denoted by $\mathcal{O}_\mu(M, G)$.

Let $\mathcal{M} = (M, \rho)$ be a finite metric space, and G be a tree connecting M . As in the case of tours, for each multitour $\pi \in \mathcal{O}_\mu(M, G)$ we put

$$p(\mathcal{M}, G, \pi) = \frac{1}{2k} \sum_{x \in X} \rho(x, \pi(x)).$$

Theorem 5.11 (A. Yu. Eremin) *For an arbitrary finite metric space $\mathcal{M} = (M, \rho)$ and an arbitrary tree G joining M , the weight of minimal parametric generalized filling can be calculated as follows*

$$\text{mpf}_-(\mathcal{M}, G) = \max\{p(\mathcal{M}, G, \pi) \mid \pi \in \mathcal{O}_\mu(M, G)\}.$$

The weight of minimal filling can be calculated as follows

$$\text{mf}(\mathcal{M}) = \text{mf}_-(\mathcal{M}) = \min_G \max\{p(\mathcal{M}, G, \pi) \mid \pi \in \mathcal{O}_\mu(M, G)\},$$

where minimum is taken over all binary trees G connecting M .

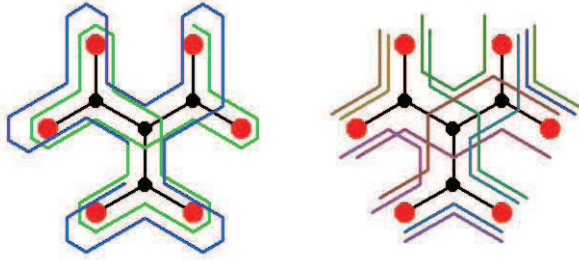


Figure 23: A part of a multitour with multiplicity 2 (left), and the irreducible boundary paths the multitour consists from (right). The multitour starts as a green polygonal line and becomes blue when multiplicity of edges becomes more than 2.

5.6 Minimal Fillings for Generic Metric Spaces

Theorem 5.11 gives an opportunity to get several interesting corollaries. To formulate one of them, we need to define what is a “generic” metric space. Notice that the set of all metric spaces consisting of n points can be naturally identified with a convex cone in $\mathbb{R}^{n(n-1)/2}$ (it suffices to enumerate the set of all two-elements subsets of these spaces and assign to each such space the vector of the distances between the pairs of points). This representation gives us an opportunity to speak about topological properties of families of metric spaces consisting of a fixed number of points.

We say, that some property holds for a *generic metric space*, if for any n this property is valid for an everywhere dense set of n -point metric spaces.

The following result can be found in [26].

Corollary 5.12 (A. Yu. Eremin) *Each general finite metric space has a minimal filling which is a nondegenerate binary tree.*

5.7 Additive Spaces and Minimal Fillings

The additive spaces are very popular in bioinformatics, playing an important role in evolution theory and, more general, in an hierarchy modeling. Recall that a finite metric space $\mathcal{M} = (M, \rho)$ is called *additive*, if M can be joined by a weighted tree $\mathcal{G} = (G, \omega)$ such that ρ coincides with the restriction of d_ω onto M . The tree \mathcal{G} in this case is called a *generating tree* for the space \mathcal{M} .

Not any metric space is additive. An additivity criterion can be stated in terms of so-called *4 points rule*: for any four points p_i, p_j, p_k, p_l , the values $\rho(p_i, p_j) + \rho(p_k, p_l)$, $\rho(p_i, p_k) + \rho(p_j, p_l)$, $\rho(p_i, p_l) + \rho(p_j, p_k)$ are the lengths of sides of an isosceles triangle whose base does not exceed its other sides.

Theorem 5.13 ([28], [29], [30], [31]) *A metric space is additive, if and only if it satisfies the 4 points rule. In the class of non-degenerate weighted trees, the generating tree of an additive metric space is unique.*

The next criterion solves completely the minimal filling problem for additive metric spaces.

Theorem 5.14 *Minimal fillings of an additive metric space are exactly its generating trees.*

The next additivity criterion is obtained by O. V. Rubleva, a student of Mechanical and Mathematical Faculty of Moscow State University, see [32].

Assertion 5.15 (O. V. Rubleva) *The weight of a minimal filling of a finite metric space is equal to the half-perimeter of this space, if and only if this space is additive.*

In the scope of Assertion 5.15, we conjectured that if there exists a tree connecting a metric space such that all the corresponding half-perimeters are equal to each other, then the space is additive. It turns out that it is not true. Z. N. Ovsyannikov suggested to consider a wider class of spaces, so called pseudo-additive spaces, for which our conjecture becomes true, see [27].

A finite metric space $\mathcal{M} = (M, \rho)$ is said to be *pseudo-additive*, if the metric ρ coincides with d_ω for a generalized weighted tree (G, ω) (which is also called *generating*), where the weight function ω can take arbitrary (not necessary nonnegative) values. Z. N. Ovsyannikov shows that these spaces can be described in terms of so-called *weak 4-points rule*: for any four points p_i, p_j, p_k, p_l , the values $\rho(p_i, p_j) + \rho(p_k, p_l)$, $\rho(p_i, p_k) + \rho(p_j, p_l)$, $\rho(p_i, p_l) + \rho(p_j, p_k)$ are the lengths of sides of an isosceles triangle. The generating tree is also unique in the class of non-degenerate trees. Moreover, the following result is valid, see [27].

Theorem 5.16 (Z. N. Ovsyannikov) *Let $\mathcal{M} = (M, \rho)$ be a finite metric space. Then the following statements are equivalent.*

- *There exist a tree G such that M coincides with the set of degree 1 vertices of G and all the half-perimeters $p(M, G, \pi)$ of M corresponding to the tours around G are equal to each other.*
- *The space \mathcal{M} is pseudo-additive.*

Moreover, the tree G in this case is a generating tree for the space \mathcal{M} .

It would be interesting to see, what role these pseudo-additive spaces could play in applications.

5.8 Examples of Minimal Fillings

Now let us give several examples of minimal filling and demonstrate how to use the technique elaborated above.

5.8.1 Triangle

Let $\mathcal{M} = (M, \rho)$ consist of three points p_1, p_2 , and p_3 . Put $\rho_{ij} = \rho(p_i, p_j)$. Consider the tree $G = (V, E)$ with $V = M \cup \{v\}$ and $E = \{vp_i\}_{i=1}^3$. Define the weight function ω on E by the following formula:

$$\omega(e_i) = \frac{\rho_{ij} + \rho_{ik} - \rho_{jk}}{2},$$

where $\{i, j, k\} = \{1, 2, 3\}$. Notice that d_ω restricted onto M coincides with ρ . Therefore, \mathcal{M} is an additive space, $\mathcal{G} = (G, \omega)$ is a generating tree for \mathcal{M} , and, due to Theorem 5.14, \mathcal{G} is a minimal filling of \mathcal{M} .

Recall that the value $(\rho_{ij} + \rho_{ik} - \rho_{jk})/2$ is called by the *Gromov product* $(p_j, p_k)_{p_i}$ of the points p_j and p_k of the space \mathcal{M} with respect to the point p_i , see [33].

5.8.2 Regular Simplex

Let all the distances in the metric space \mathcal{M} are the same and are equal to d , i.e. \mathcal{M} is a regular simplex. Then the weighted tree $\mathcal{G} = (G, \omega)$, $G = (V, E)$, with the vertex set $V = M \cup \{v\}$ and edges vm , $m \in M$, of the weight $d/2$ is generating for \mathcal{M} . Therefore, the space \mathcal{M} is additive, and, due to Theorem 5.14, \mathcal{G} is its unique nondegenerate minimal filling. If n is the number of points in M , then the weight of the minimal filling is equal to $dn/2$.

5.8.3 Star

If a minimal filling $\mathcal{G} = (G, \omega)$ of a space $\mathcal{M} = (M, \rho)$ is a star whose single interior vertex v is joined with each point $p_i \in M$, $1 \leq i \leq n$, $n \geq 3$, then the metric space \mathcal{M} is additive [3]. In this case the weights of edges can be calculated easily. Indeed, put $e_i = vp_i$. Since a subspace of an additive space is additive itself, then we can use the results for three-points additive space, see above. So, we have $\omega(e_i) = (p_j, p_k)_{p_i}$, where p_i, p_j , and p_k are arbitrary distinct boundary vertices.

5.8.4 Mustaches of Degree more than 2

Let $G = (V, E)$ be an arbitrary tree, and $v \in V$ be an interior vertex of degree $(k+1) \geq 3$ adjacent with k vertices w_1, \dots, w_k from ∂G . Then the set of the vertices $\{w_1, \dots, w_k\}$, and also the set of the edges $\{vw_1, \dots, vw_k\}$, are referred as *mustaches*. The number k is called by the *degree*, and the vertex v is called by the *common vertex of the mustaches*. An edge incident to v and not belonging to $\{vw_1, \dots, vw_k\}$ is called the *root edge* of the mustaches under consideration.

As it is shown in [3], any mustaches of a minimal filling of a metric space forms an additive subspace. If the degree of such mustaches is more than 2, then we can calculate the weights of all the edges containing in the mustaches just in the same way as in the case of a star.

5.8.5 Four-Points Spaces

Here we give a complete description of minimal fillings for four-points spaces, see details in [3].

Proposition 5.17 *Let $M = \{p_1, p_2, p_3, p_4\}$, and ρ be an arbitrary metric on M . Put $\rho_{ij} = \rho(p_i, p_j)$. Then the weight of a minimal filling $\mathcal{G} = (G, \omega)$ of the space $\mathcal{M} = (M, \rho)$ is given by the following formula*

$$\frac{1}{2}(\min\{\rho_{12} + \rho_{34}, \rho_{13} + \rho_{24}, \rho_{14} + \rho_{23}\} + \max\{\rho_{12} + \rho_{34}, \rho_{13} + \rho_{24}, \rho_{14} + \rho_{23}\}).$$

If the minimum in this formula is equal to $\rho_{ij} + \rho_{rs}$, then the type of minimal filling is the binary tree with the mustaches $\{p_i, p_j\}$ and $\{p_r, p_s\}$.

We apply the obtained result to the vertex set of a planar convex quadrangle.

Corollary 5.18 *Let M be the vertex set of a convex quadrangle $p_1p_2p_3p_4 \subset \mathbb{R}^2$ and $\rho(p_i, p_j) = \|p_i - p_j\|$. The weight of a minimal filling of the space (M, ρ) is equal to*

$$\frac{1}{2} \min(\rho_{12} + \rho_{34}, \rho_{14} + \rho_{23}) + \frac{\rho_{13} + \rho_{24}}{2}.$$

The topology of minimal filling is a binary tree with mustaches corresponding to opposite sides of the less total length.

5.9 Ratios

The Steiner ratio is discussed in Section 4. Here we define some other ratios based on minimal fillings, which could be more available for calculating, and which could be useful to calculate the Steiner ratio, as we hope.

5.10 Steiner–Gromov Ratio

For convenience, the sets consisting of more than a single point are referred as *nontrivial*. Let $\mathcal{X} = (X, \rho)$ be an arbitrary metric space, and let $M \subset X$ be some finite subset. Recall that by $\text{mst}(M, \rho)$ we denote the length of minimal spanning tree of the space (M, ρ) . Further, for nontrivial M , we define the value

$$\text{sgr}(M) = \text{mf}(M, \rho) / \text{mst}(M, \rho)$$

and call it the *Steiner–Gromov ratio* of the subset M . The value $\inf \text{sgr}(M)$, where the infimum is taken over all nontrivial finite subsets of \mathcal{X} , consisting of at most n vertices is denoted by $\text{sgr}_n(\mathcal{X})$ and is called the *degree n Steiner–Gromov ratio of the space \mathcal{X}* . At last, the value $\inf \text{sgr}_n(\mathcal{X})$, where the infimum is taken over all positive integers $n > 1$ is called the *Steiner–Gromov ratio of the space \mathcal{X}* and is denoted by $\text{sgr}(\mathcal{X})$, or by $\text{sgr}(X)$, if it is clear what particular metric on X is considered. Notice that $\text{sgr}_n(\mathcal{X})$ is a non-increasing function on n . Besides, it is easy to see that $\text{sgr}_2(\mathcal{X}) = 1$ and $\text{sgr}_3(\mathcal{X}) = 3/4$ for any nontrivial metric space \mathcal{X} .

Assertion 5.19 *The Steiner–Gromov ratio of an arbitrary metric space is not less than $1/2$. There exist metric spaces whose Steiner–Gromov ratio equals to $1/2$.*

Recently, A. Pakhomova, a student of Mechanical and Mathematical Faculty of Moscow State University, obtained an exact general estimate for the degree n Steiner–Gromov ratio, see [35].

Assertion 5.20 (A. Pakhomova) *For any metric space \mathcal{X} the estimate*

$$\text{sgr}_n(\mathcal{X}) \geq \frac{n}{2n-2}$$

is valid. Moreover, this estimate is exact, i.e. for any $n \geq 3$ there exists a metric space \mathcal{X}_n such that $\text{sgr}_n(\mathcal{X}_n) = n/(2n-2)$.

Also recently, Z. Ovsyannikov [36] investigated the metric space of all compact subsets of Euclidean plane endowed with Hausdorff metric.

Assertion 5.21 (Z. Ovsyannikov) *The Steiner ratio and the Steiner–Gromov ratio of the metric space of all compact subsets of Euclidean plane endowed with Hausdorff metric are equal to $1/2$.*

5.11 Steiner Subratio

Let $\mathcal{X} = (X, \rho)$ be an arbitrary metric space, and let $M \subset \mathcal{X}$ be some its finite subset. Recall that by $\text{smt}(M, \rho)$ we denote the length of Steiner minimal tree joining M . Further, for nontrivial subsets M , we define the value

$$\text{ssr}(M) = \text{mf}(M, \rho) / \text{smt}(M, \rho)$$

and call it by the *Steiner subratio* of the set M . The value $\inf \text{ssr}(M)$, where infimum is taken over all nontrivial finite subsets of \mathcal{X} consisting of at most $n > 1$ points, is denoted

by $\text{ssr}_n(\mathcal{X})$ and is called the *degree n Steiner subratio of the space \mathcal{X}* . At last, the value $\inf \text{ssr}_n(\mathcal{X})$, where the infimum is taken over all positive integers $n > 1$, is called the *Steiner subratio of the space \mathcal{X}* and is denoted by $\text{ssr}(\mathcal{X})$, or by $\text{ssr}(X)$, if it is clear what particular metric on X is considered. Notice that $\text{ssr}_n(\mathcal{X})$ is a nonincreasing function on n . Besides, it is easy to see that $\text{ssr}_2(\mathcal{X}) = 1$ for any nontrivial metric space \mathcal{X} .

Proposition 5.22 $\text{ssr}_3(\mathbb{R}^n) = \sqrt{3}/2$.

The next result is obtained by E. I. Filonenko, a student of Mechanical and Mathematical Department of Moscow State University, see [34].

Proposition 5.23 (E. I. Filonenko) $\text{ssr}_4(\mathbb{R}^2) = \sqrt{3}/2$.

Conjecture 5.24 *The Steiner subratio of the Euclidean plane is achieved at the regular triangle and, hence, is equal to $\sqrt{3}/2$.*

Recently, A. Pakhomova obtained an exact general estimate for the degree n Steiner subratio, see [35].

Proposition 5.25 (A. Pakhomova) *For any metric space $\{\mathcal{X}\}$ the estimate*

$$\text{ssr}_n(\mathcal{X}) \geq \frac{n}{2n-2}$$

is valid. Moreover, this estimate is exact, i.e. for any $n \geq 3$ there exists a metric space \mathcal{X}_n such that $\text{ssr}_n(\mathcal{X}_n) = n/(2n-2)$.

Also recently, Z. Ovsyannikov [36] investigated the metric space of all compact subsets of Euclidean plane endowed with Hausdorff metric.

Proposition 5.26 (Z. Ovsyannikov) *Let \mathcal{C} be the metric space of all compact subsets of Euclidean plane endowed with Hausdorff metric. Then $\text{ssr}_3(\mathcal{C}) = 3/4$ and $\text{ssr}_4(\mathcal{C}) = 2/3$.*

References

- [1] A. O. Ivanov, and A. A. Tuzhilin. Branching Solutions to One-Dimensional Variational Problems. Singapore, New Jersey, London, Hong Kong, 2001. 342 p.
- [2] A. O. Ivanov, and A. A. Tuzhilin. Extreme Networks Theory. Moscow, Izhevsk, 2003. 406 p. [In Russian.]
- [3] A. O. Ivanov, and A. A. Tuzhilin. One-dimensional Gromov minimal filling problem. // Sbornik: Mathematics. 2012. V. 203, No 5. P. 677-726 [Matem. sb. 2012. T. 203, No 5. S. 65-118].
- [4] V. A. Emelichev, at al. Lectures on Graph Theory. Moscow, 1990. 384 p. [In Russian.]
- [5] B. Chazelle. The soft heap: an approximate priority queue with optimal error rate. // Journal of the Association for Computing Machinery. 2000. V. 47, No 6. P. 1012-1027.
- [6] A. O. Ivanov and A. A. Tuzhilin. The geometry of inner spanning trees for planar polygons. // Izvestiya: Mathematics. 2012. V. 76, No 2. P. 215-244 [Izv. RAN. Ser. Matem. 2012. T. 76, No 2. S. 3-36].

- [7] *A. O. Ivanov and A. A. Tuzhilin.* The Steiner problem in the plane or in plane minimal nets. // Mathematics of the USSR-Sbornik. 1993. V. 74, No 2. P. 555–582 [Matem. Sb. 1991. T. 182, No 12. S. 1813–1844].
- [8] *F. K. Hwang.* A linear time algorithm for full Steiner trees. // Oper. Res. Letter. 1986. V. 5. P. 235–237.
- [9] *Z. A. Melzak.* On the problem of Steiner. // Canad. Math. Bull. 1960. V. 4. P. 143–148.
- [10] *A. O. Ivanov and A. A. Tuzhilin.* Solution of the Steiner problem for convex boundaries. // Russian Mathematical Surveys. 1990. V. 45, No 2. P. 214–215 [Uspekhi Matem. Nauk. 1990. T. 45, No 2(272). S. 207–208].
- [11] *E. N. Gilbert and H. O. Pollak.* “Steiner minimal trees.” // SIAM J. Appl. Math. 1968. V. 16, No 1. P. 1–29.
- [12] *D. Cieslik.* The Steiner ratio. Springer, 2001. 256 p.
- [13] *D. Z. Du, F. K. Hwang.* The Steiner ratio conjecture of Gilbert–Pollak is true. // Proc. Nat. Acad. Sci. 1990. V. 87. P. 9464–9466.
- [14] *D. Z. Du, F. K. Hwang.* A proof of Gilbert–Pollak Conjecture on the Steiner ratio. // Algorithmica. 1992. V. 7. P. 121–135.
- [15] *A. O. Ivanov and A. A. Tuzhilin.* The Steiner Ratio GilbertPollak Conjecture Is Still Open. Clarification Statement. // Algorithmica. 2012. V. 62. No 1–2. P. 630–632.
- [16] *P. O. de Wet.* Geometric Steiner Minimal Trees. PhD Thesis. UNISA, Pretoria, 2008.
- [17] *D. Z. Du and W. D. Smith.* Disproofs of Generalized Gilbert–Pollack conjecture on the Steiner ratio in three and more dimensions. // Combin. Theory. 1996. V. 74. Ser. A. P. 115–130.
- [18] *W. D. Smith and J. M. Smith.* On the Steiner ratio in 3-Space. // J. of Comb. Theory. 1995. V. 65. Ser. A. P. 301–322.
- [19] *R. L. Graham and F. K. Hwang.* A remark on Steiner minimal trees. // Bull. of the Inst. of Math. Ac. Sinica. 1976. V. 4. P. 177–182.
- [20] *N. Innami and B. H. Kim.* Steiner ratio for Hyperbolic surfaces. // Proc. Japan Acad. 2006. V. 82. Ser. A. No 6. P. 77–79.
- [21] *D. Cieslik, A. Ivanov and A. Tuzhilin.* Steiner Ratio for Manifolds. // Mat. Zametki. 2003. V. 74. No 3. P. 387–395 [Mathematical Notes. 2003. V. 74. No 3. P. 367–374].
- [22] *M. Gromov.* Filling Riemannian manifolds. // J. Diff. Geom. 1983. V. 18. No 1. P. 1–147.
- [23] *A. O. Ivanov, A. A. Tuzhilin.* Steiner Ratio. The State of the Art. // Matemat. Voprosy Kibern. 2002. V. 11. P. 27–48 [In Russian].
- [24] *P. A. Borodin.* An example of nonexistence of a steiner point in a Banach space. // Matem. Zametki. 2010. V. 87. No 4. P. 514–518 [Mathematical Notes. 2010. V. 87. No 4. P. 485–488].
- [25] *A. O. Ivanov, Z. N. Ovsyannikov, N. P. Strelkova, A. A. Tuzhilin.* One-dimensional minimal fillings with edges of negative weight. // Vestnik MGU, ser. Mat., Mekh. 2012. No 5. P. 3–8.

- [26] *A. Yu. Eremin*. Formula calculating the weight of minimal filling. // *Matem. Sbornik*. 2012. [To appear].
- [27] *Z. N. Ovsyannikov*. Pseudo-additive metric spaces and minimal fillings. // *Vestnik MGU*. 2013. [To appear].
- [28] *K. A. Zareckii*. Constructing a tree on the basis of a set of distances between the hanging vertices. // *Uspehi Mat. Nauk*. 1965. V. 20. No 6. P. 90-92. [In Russian].
- [29] *J. M. S. Simões-Pereira*. A note on the tree realizability of a distance matrix. // *J. Combinatorial Th.* 1969. V. 6. P. 303–310.
- [30] *E. A. Smolenskij*. About a Linear Denotation of Graphs. // *J. vych. matem. i matem. phys.* 1962. V. 2. No 2. P. 371–372. [In Russian].
- [31] *S. L. Hakimi, S. S. Yau*. Distane matrix of a graph and its realizability. // *Quart. Appl. Math.* 1975. V. 12. P. 305–317.
- [32] *O. V. Rubleva*. Additivity Criterion for finite metric spaces and minimal fillings. // *Vestnik MGU, ser. matem., mekh.* 2012. No 2. P. 8–11.
- [33] *M. Gromov*. Hyperbolic groups. // in book: *S. M. Gersten, ed.* *Essays in Group Theory*. 1987. Springer.
- [34] *E. I. Filonenko*. Degree 4 Steiner subratio of Euclidean plane. // *Vestnik MGU, ser. matem., mekh.* 2013. To appear.
- [35] *A. S. Pakhomova*. Estimates for Steiner–Gromov ratio and Steiner subratio. // *Vestnik MGU, ser. matem., mekh.* 2013. To appear.
- [36] *Z. N. Ovsyannikov*. Steiner ratio, Steiner–Gromov ratio and Steiner subratio for the metric space of all compact subsets of the Euclidean space with Hausdorff metric. // *Vestnik MGU, ser. matem., mekh.* 2013. To appear.